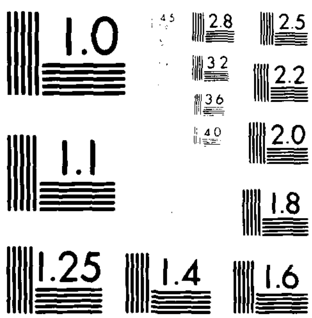
AD-A085 929 BROWN UNIV PROVIDENCE RI LEFSCHETZ CENTER FOR DYNAMI--ETC F/G 12/2
PARAMETER ESTIMATION AND IDENTIFICATION FOR SYSTEMS WITH DELAYS--ETC(U)
NOV 79 H T BANKS, J A BURNS, E M CLIFF DAAG29-79-C-0161

UNCLASSIFIED AFOSR-TR-80-0403 NL

1 OF 1
AD
A085929

END
DATE
FILMED
8-80
DTIC



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

AFOSR-TR- 80-0403

PARAMETER ESTIMATION AND IDENTIFICATION

FOR

SYSTEMS WITH DELAYS

3 LEVEL II

ADA085929

November 1979

88

9 Interim rept.

H. T. Banks*

Lefschetz Center for Dynamical Systems
Division of Applied Mathematics
Brown University
Providence, Rhode Island 02912

J. A. Burns+

Department of Mathematics
Virginia Polytechnic Institute and State University
Blacksburg, Virginia 24061

E. M. Cliff

Aerospace and Ocean Engineering Department
Virginia Polytechnic Institute and State University
Blacksburg, Virginia 24061

*This research was supported in part by the National Science Foundation under NSF-MCS79-05774, in part by the Air Force Office of Scientific Research under AF-AFOSR 76-3092A and in part by the U.S. Army Research Office under DAAG29-79-C-0161.

VAFOSR-76-3092

+This research was supported in part by the U.S. Army Research Office under ARO-DAAG-29-78-G-0125.

PARAMETER ESTIMATION AND IDENTIFICATION
FOR
SYSTEMS WITH DELAYS

H. T. Banks, J. A. Burns and E. M. Cliff

Abstract

Parameter identification problems for delay systems motivated by examples from aerodynamics and biochemistry are considered. The problem of estimation of the delays is included. Using approximation results from semigroup theory, a class of theoretical approximation schemes is developed and two specific cases ("averaging" and "spline" methods) are shown to be included in this treatment. Convergence results, error estimates, and a sample of numerical findings are given.

1. Introduction

The estimation of parameters in dynamical systems is an important scientific problem on which a number of contributions have been made in the engineering and mathematical literature (e.g. see [1], [23]). However, for systems with delays, very little on identification is found in the engineering literature and essentially no theoretical convergence results are available for algorithms dealing with estimation of the delays themselves. One obvious difficulty (from both a practical and theoretical viewpoint) with such procedures is that solutions of delay systems are not in general differentiable with respect to the delays and thus many common (e.g. least squares gradient, maximum likelihood estimator, etc.) identification techniques are not directly applicable.

In this paper we discuss a class of methods based on general approximation techniques for systems with delays. These approximation ideas have been considered earlier in the context of optimal control problems ([3], [4], [5], [6], [7], [9], [12], [18]) where they have proved quite useful. The use of such approximation ideas in connection with parameter estimation procedures was apparently first suggested in [11] and some preliminary theoretical results were stated in [7] and [13]. However, our presentation here is the first (to our knowledge) rigorous treatment of general theoretical aspects of these ideas.

While we do in section 7 below give a small sample of related numerical findings, the primary purpose of this manuscript is to present a theoretical foundation for the schemes we propose. A much more extensive discussion and a wider selection of numerical examples is presented in [8]. Our sample of numerical results in section 7 is included here mainly to indicate that the procedures based on the schemes discussed actually are feasible.

The approximation ideas developed earlier in [5] and employed here are based on approximation results (the so-called Trotter-Kato theorem) from linear semigroup theory. In section 2 we formulate a class of identification problems for delay systems and show that they can be reformulated in an abstract setting so as to make use of the semigroup approximation theorem. A version of the Trotter-Kato results needed is given in section 3, while in section 4 we show how to use this theorem to insure convergence for a class of identification schemes. We turn to the detailed development of particular schemes based on "Averaging" (see [5]) and "Spline" (see [10]) approximations in the subsequent two sections. Finally, a brief indication of numerical findings for these two particular schemes is given in section 7.

Notation used throughout the paper is completely standard. For example, $L_p^m(a,b) = L_p([a,b],R^m)$ denotes the usual Lebesgue spaces of R^m-valued "functions" on [a,b] whose components are integrable when raised to the pth power. When $m = 1$, we shall suppress its appearance in the notation. $L_{p,loc}$ denotes the usual "locally" integrable function spaces. We shall use the symbol $|\cdot|$ to denote the norm of an element without distinguishing between different norms if the intended meaning is clear from the context. The space of functions with j continuous derivatives is denoted by $C^j(a,b)$. We shall also make use of the Sobolev spaces $W_p^{(j)}(a,b) = W_p^{(j)}([a,b],R^n)$ of R^n-valued absolutely continuous functions possessing j-1 absolutely continuous derivatives and jth derivatives that are in L_p.

In the remaining paragraphs of this introductory section, we turn to a discussion of examples which motivate the theoretical questions that are the focus of our attention in this paper.

1.1. Tubular reactor columns and delay system identification and control problems

Packed bed tubular enzyme reactors are very important in many areas of industrial and biological applications (potential uses involve purification or classification of fruit juices, proteolytic treatment of beer, synthesis of essential amino acids, enzymatic biosynthesis - i.e., synthesis of antibiotics and steroids, etc.). These are column reactors (as depicted in Figure 1.1) containing enzyme pellets (i.e. pellets in which an enzyme is insolubly bound), the enzyme being specific for a substrate S which is passed through the column. The substrate diffuses into the pellets where the enzyme catalyzes a reaction resulting in the product P.

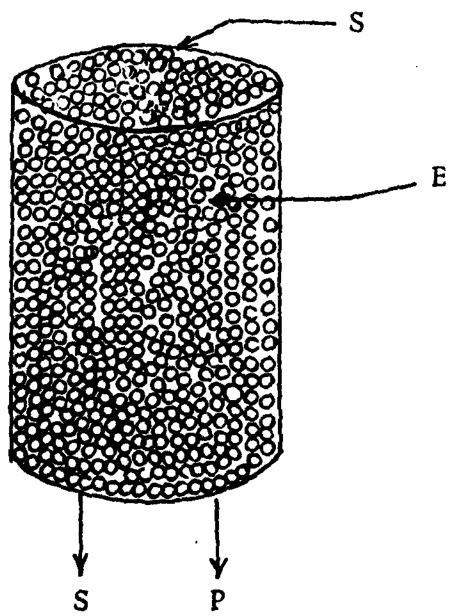


FIGURE 1.1

We thus have enzymatically active particles or pellets in a convective flow region. Any model should embody important features of the system including (i) enzyme catalyzed reaction, (ii) metabolite (S or P) diffusion into, out of, and inside of the pellets, and (iii) metabolite convection (and possibly diffusion) in the flow region in the column exterior to the pellets. Extensive studies for both plug-flow (PF) models and diffusion-convection-reaction (DCR) models for the phenomena involved have been reported in the literature [15], [16]. These models can be formulated from first principles using transport equations of the form

$$\frac{\partial s}{\partial t} + c\,\frac{\partial s}{\partial x} = D_1\,\frac{\partial^2 s}{\partial x^2} + D_2\,\frac{\partial^2 s}{\partial y^2} + V$$

where D_1, D_2 are diffusion coefficients, c is the convective flow velocity, x is the column axial direction, y the perpendicular direction (in a 2-dimensional model), and V is a nonlinear reaction velocity approximation (e.g., $V = -\rho\,\frac{s}{1+s}$). The column in this case is approximated by a two-compartment (pellet phase and liquid phase) model as shown in Figure 1.2.

PF models incorporate assumptions that one may ignore diffusion in both the pellet and solution (flow) regions. Careful investigation of these models reveal that they are of limited use in actual applications since it is found that certain kinetic <u>constants</u> must actually be allowed to vary (in an unpredictable manner) with the flow velocity in order to fit the models to experimental data. On the other hand, the DCR models were found to perform quite adequately when compared with the data. The main difficulty in employing the

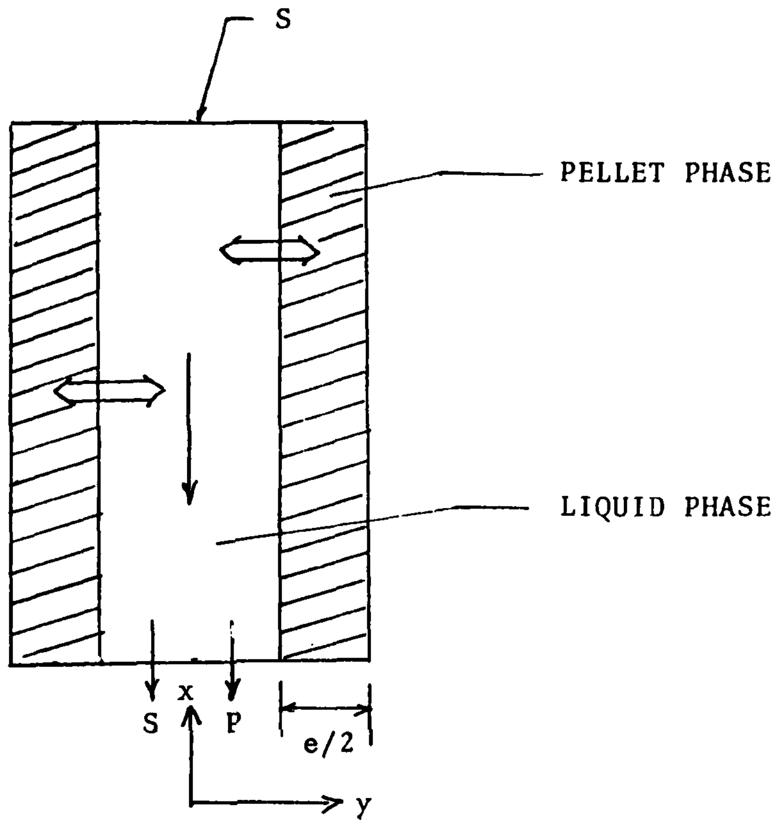


FIGURE 1.2

DCR models involves the rather lengthy calculations that must be made in carrying out identification and control procedures with these models. It is, therefore, desirable to have a model which in complexity and accuracy (hopefully similar to the PF models with respect to the latter) is somewhere between the PF and DCR models and for which efficient numerical procedures are available.

A candidate for such a model has been proposed by J.P. Kernevez and his colleagues at Université de Technologie de Compiègne. It consists of n functional compartments for the column, each containing two subcompartments, one representing the pellet phase and the other the liquid phase. The two subcompartments in each compartment are connected by diffusion while the main compartments are connected via unidirectional transport (convective flow) between the liquid phase subcompartments (see Figure 1.3).

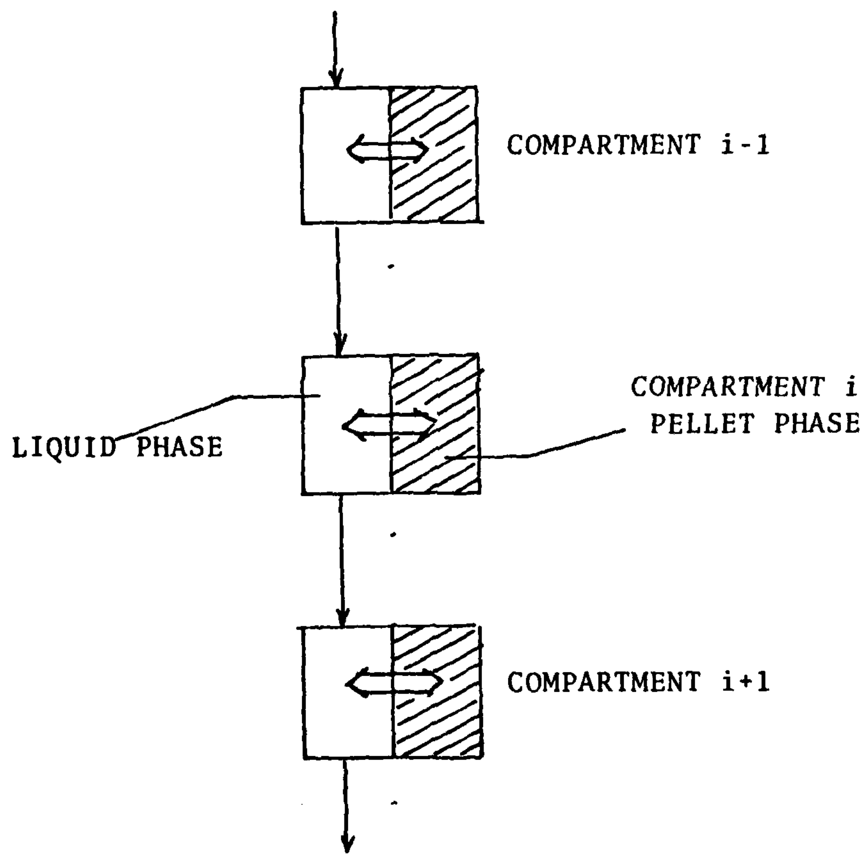


FIGURE 1.3

Defining variables as follows (all are scaled and dimensionless):

$r_i(t)$ = substrate concentration in liquid phase in compartment i at time t,

$s_i(t)$ = substrate concentration in pellet phase in compartment i at time t,

$p_i(t)$ = product concentration in liquid phase in compartment i at time t,

$q_i(t)$ = product concentration in pellet phase in compartment i at time t,

one can write mass balance equations to obtain a model

$$\frac{dr_1}{dt} = -\alpha r_1(t) - \beta\{r_1(t) - s_1(t-\tau_1)\} + u(t)$$

$$\frac{dr_i}{dt} = \alpha\{r_{i-1}(t-\tau) - r_i(t)\} - \beta\{r_i(t) - s_i(t-\tau_1)\}, \quad i > 1,$$

$$\frac{ds_i}{dt} = -\rho F(s_i(t)) + \nu\beta\{r_i(t) - s_i(t-\tau_1)\}, \quad i \geq 1,$$

$$\frac{dp_1}{dt} = -\alpha p_1(t) - \tilde{\beta}\{p_1(t) - q_1(t-\tau_2)\}$$

$$\frac{dp_i}{dt} = \alpha\{p_{i-1}(t-\tau) - p_i(t)\} - \tilde{\beta}\{p_i(t) - q_i(t-\tau_2)\}, \quad i > 1,$$

$$\frac{dq_i}{dt} = \rho F(s_i(t)) + \nu\tilde{\beta}\{p_i(t) - q_i(t-\tau_2)\}, \quad i \geq 1.$$

Here F is a nonlinear reaction velocity term, the delays τ, τ_1, τ_2 are transport times between compartments $i-1$ and i, between pellet interior and liquid region for substrate, and between pellet interior and liquid region for product, respectively. The term u represents input of substrate to the liquid subcompartment of compartment 1. The parameters $\alpha, \beta, \rho, \nu, \tilde{\beta}$ are all related to biochemical and physical constants for the column configuration. For example, $\beta \tilde{=} ND_s\Sigma/eV$ where N is the "apparent" number of pellets per compartment, V = volume of liquid per compartment, D_s = coefficient of diffusion for S within the pellet region, e = "thickness" of the model pellet region, and Σ = effective surface area per pellet.

Using data collected from a number of specific experiments performed with tracer, product, and substrate inputs, one wishes to determine values of $\tau, \tau_1, \tau_2, \rho, \beta, \tilde{\beta}$ so that the model describes accurately the operation of the column. Once this is done, the model then must be used to design (optimal) control procedures for the column.

We thus have classical identification and control problems for systems (let $x_i = (r_i, s_i, p_i, q_i)^T$)

$$\dot{x}_i(t) = A_0(\gamma)x_i(t) + A_1(\gamma)x_i(t-\tau_1) + A_2(\gamma)x_i(t-\tau_2)$$

$$+ f(\gamma, x_{i-1}(t-\tau), x_i(t), u(t))$$

where the delays τ, τ_1, τ_2 and the vector parameter γ (involving only coefficients) are to be identified.

1.2. Identification problem for hereditary systems in unsteady aerodynamics

We consider next an interesting class of identification problems which arise in the study of unsteady aerodynamics (see [2]). Consider a thin, flat airfoil mounted on springs as shown in Figure 1.4 in a region where we have fluid (air) flow with undisturbed stream velocity U (in the x-direction). Flow around the airfoil is disturbed and we assume it has velocity $\bar{q} = (u,w)$. Laws of conservation of mass and momentum lead to a system of partial differential equations for the fluid velocity components u and w. Assuming incompressible flow we have the continuity equation $\nabla \cdot \bar{q} = 0$. Elementary hydrodynamics also yield that curl $(\bar{q}) = 0$ from which we deduce the existence of a velocity potential φ so that $\bar{q} = \nabla\varphi$. The equation of continuity then becomes $\Delta\varphi = 0$. We restrict our considerations to small motions of the airfoil so that a linearized theory may be adopted. We assume that φ is given by

$$\varphi(x,z,t) = \tilde{\varphi}(x,z,t) + Ux$$

where $\tilde{\varphi}$ is a disturbance potential. It follows that $\tilde{\varphi}$ must satisfy

$$\Delta\tilde{\varphi} = 0. \tag{1.1}$$

In addition one has the (flow tangency) boundary conditions

$$(1.2) \qquad \frac{\partial\tilde{\varphi}}{\partial z}(x,0,t) = w(x,0,t) = w_a(x,t) \qquad -1 \le x \le 1$$

where w_a is a given function describing the motion of the airfoil. We here assume that the airfoil is a thin plate located at $z = 0$, $-1 \le x \le 1$ as depicted in Figure 1. 5.

After arguments involving a conformal mapping of the airfoil into the unit circle and the introduction of sources (elementary flows along radial lines) and vortices (elementary flows along concentric circles), one finds that a solution of (1.1),(1.2) for the disturbance potential $\tilde{\varphi}$ consists of an appropriate collection of sources distributed along the airfoil and any weighted combination of "compatible" vortex pairs. A "compatible" pair consists of one vortex on the airfoil at $r = r_1 < 1$ and an oppositely rotating one at $r = \frac{1}{r_1} > 1$. Compatible pairs induce a flow with finite angular momentum and with fluid velocity that is tangent to the airfoil. The required distribution of sources is uniquely defined by the airfoil motion ($w_a(x,t)$ in equation (1.2)) but the distribution of vortices in the wake given by a density function $\gamma_w(\xi,t)$ is as yet unknown. In lieu of γ_w we introduce a new function Γ termed the circulation. For brevity we shall "<u>define</u>" Γ by

$$(1.3) \qquad \dot{\Gamma}\left(t - \frac{(\xi-1)}{U}\right) = -\gamma_w(\xi,t)$$

with the boundary condition $\Gamma(-\infty) = 0$. This relationship reveals that vorticity in the wake at time t and position ξ was produced by a change in the circulation at an earlier time, i.e. an hereditary phenomenon is involved. In integrated form (using

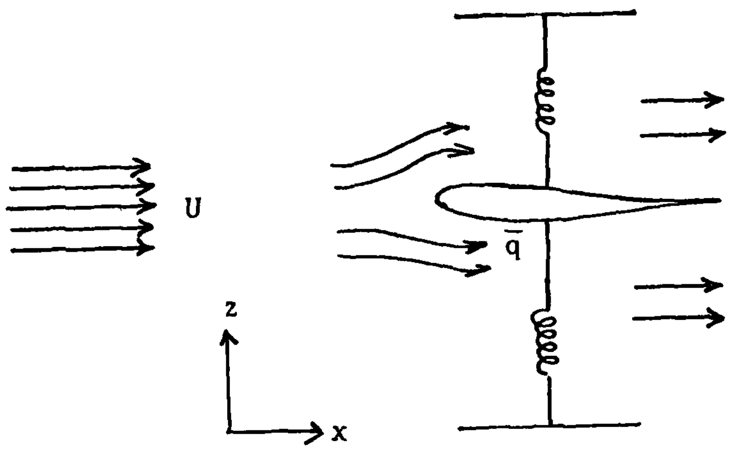


FIGURE 1.4

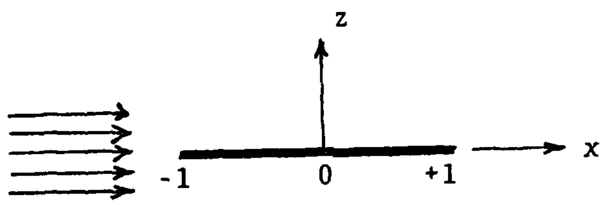


FIGURE 1.5

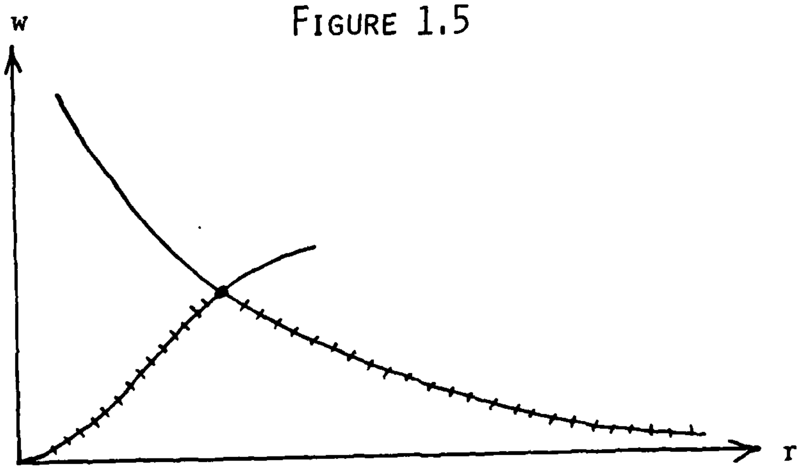


FIGURE 1.6

$\Gamma(-\infty) = 0$) this becomes

$$(1.4) \qquad \Gamma(t) - \int_1^\infty \gamma_w(\xi,t)d\xi = 0.$$

To determine Γ (or γ_w) we impose an additional hypothesis, viz. finiteness of the fluid velocity at the trailing edge of the airfoil. Mathematically this is written

$$(1.5) \qquad v(t) + \int_1^\infty \sqrt{\frac{\xi+1}{\xi-1}}\ \gamma_w(\xi,t)d\xi = 0.$$

Here v is the contribution to the velocity due to the source distribution. Subtracting (1.5) from (1.4) and using (1.3) we thus obtain

$$(1.6) \qquad \Gamma(t) = v(t) + \int_1^\infty f(\xi)\dot{\Gamma}(t - \frac{(\xi-1)}{U})d\xi$$

where $f(\xi) = \sqrt{(\xi+1)/(\xi-1)} - 1$. This finally is our model equation (see [28], p. 292) that is the basis of hereditary models in unsteady aerodynamics.

A simple change of variables $\xi = 1 - \sigma$ in the integral in (1.6) yields the equation

$$(1.7) \qquad \Gamma(t) = v(t) + \int_{-\infty}^0 \tilde{f}(\sigma)\dot{\Gamma}(t + \frac{\sigma}{U})d\sigma$$

where $\tilde{f}(\sigma) \equiv f(1-\sigma)$. This is essentially a neutral functional differential equation with infinite memory. Among the numerous approximations made in the derivation of such a model is the expression for f

$$f(\xi) = \sqrt{(\xi+1)/(\xi-1)} - 1. \tag{1.8}$$

It turns out that the transverse velocity component w exhibits a boundary layer phenomenon as sketched in Figure 1.6 where r is the horizontal distance from the trailing edge of the airfoil. Thus, the expression for f in (1.8) is valid only for $\xi >> 1$. To better approximate this phenomenon in (1.7) one might approximate $\tilde{f}$ by a function g having the form

$$g(\sigma;\alpha,\beta,\mu) = \begin{cases} \alpha\sigma & -\mu \leq \sigma \leq 0 \\ \beta\tilde{f}(\sigma) & -\infty < \sigma < -\mu \end{cases}$$

when it is understood that α, β must be chosen so that g is continuous at $\sigma = -\mu$. The model then is given by

$$\Gamma(t) = v(t) + \int_{-\infty}^{0} g(\sigma;\alpha,\beta,\mu)\dot{\Gamma}(t + \frac{\sigma}{U})d\sigma. \tag{1.9}$$

Assuming smoothness of g we formally integrate by parts the integral in (1.9) to obtain

$$\Gamma(t) = v(t) - \int_{-\infty}^{0} \dot{g}(\sigma)\Gamma(t + \frac{\sigma}{U})d\sigma$$

or letting $s = t + \frac{\sigma}{U}$ in the integral and defining $G(\xi) = \dot{g}(U\xi)$, we have

$$\Gamma(t) = v(t) - \int_{-\infty}^{t} G(s-t)\Gamma(s)ds. \tag{1.10}$$

Equation (1.10) is a retarded FDE with infinite memory which, upon differentiation, yields the more familiar form

$$\dot{\Gamma}(t) = \dot{v}(t) - G(0)\Gamma(t) + \int_{-\infty}^{t} \dot{G}(s-t)\Gamma(s)ds.$$

In practice, one often would desire to replace the integral term by a finite integral

$$\int_{-\tau}^{t} \dot{G}(s-t)\Gamma(s)ds$$

in which case one obtains an equation (taking $\Gamma(t) = x(t)$, observing that $G(0) = \dot{g}(0) = \alpha$ and identifying $\dot{v}(t) = u(t)$ as the input)

$$\text{(1.11)} \quad \dot{x}(t) = -\alpha x(t) + \int_{-\tau}^{t} U\ddot{g}(U[s-t];\alpha,\beta,\mu)x(s)ds + u(t).$$

An important identification problem then consists of making observations corresponding to an input $u(t) = \dot{v}(t)$ and using these to estimate the parameters α,β,μ, and τ so that the model yields a sufficiently accurate description of the aerodynamic phenomena under investigation.

2. The fundamental identification problem for delay systems

We consider in this paper n-vector systems of the form

$$(2.1)\qquad \dot{x}(t) = L(q)x_t + B(\alpha)u(t), \quad t \geq 0,$$

with initial data

$$(2.2)\qquad x(0) = \eta,\ x_0 = \phi, \quad (\eta,\phi) \in R^n \times L_2^n(-r,0),$$

and output

$$(2.3)\qquad y(t) = C(\alpha)x(t) + D(\alpha)u(t).$$

We make the following definitions and assumptions about the operators and parameters in (2.1)-(2.3). There exists a fixed given $r > 0$ and compact convex set $\Omega \subset R^\mu$ and we define the compact convex set $Q \subset R^{\mu+\nu}$ by $Q \equiv \Omega \times \mathscr{A}$, where

$$\mathscr{A} = \{h = (r_1,r_2,\ldots,r_\nu) \in R^\nu \mid 0 \leq r_i \leq r_{i+1} \leq r,\ i = 1,\ldots,\nu-1\}.$$

For a function x we adopt the usual notation $x_t(\theta) = x(t+\theta)$. For a given element $q = (\alpha,h)$ in the admissible parameter set Q, we define the operators $L(q)\colon L_2^n(-r,0) \to R^n$ of (2.1) by

$$(2.4)\qquad L(q)\phi = \sum_{i=0}^{\nu} A_i(\alpha)\phi(-r_i) + \int_{-r_\nu}^{0} K(\alpha,\theta)\phi(\theta)d\theta$$

where $r_0 \equiv 0$, and for each $\alpha \in \Omega$, $A_i(\alpha), B(\alpha), C(\alpha)$, and $D(\alpha)$ are

$n \times n$, $n \times m$, $k \times n$, and $k \times m$ matrices respectively. We assume that the $n \times n$ matrix valued function $\theta \to K(\alpha,\theta)$ is in $L_2(-r,0)$ and that the functions $A_i, B, C, D, K(\ ,\cdot)$ are continuous in α.

Remark 1: In (2.4) one must give the proper interpretation to point evaluations in the event ϕ is only an L_2 "function". Since in (2.1) we are interested in integrals of the system, the usual interpretation is intended here (see [10] for a more detailed discussion).

We further assume that we are given an initial data set $\mathcal{S} \subset R^n \times L_2^n(-r,0)$ that is closed, bounded and convex and we define

$$\Gamma \equiv \mathcal{S} \times Q = \mathcal{S} \times \Omega \times \mathcal{A}$$

as our admissible initial data - parameter set. Elements γ in Γ will be denoted in one of several ways throughout our discussions below:

$$\gamma = (\eta,\phi,q) = (\eta,\phi,\alpha,h) = (\eta,\phi,\alpha,r_1,\ldots,r_\nu)$$

where $q = (\alpha,h) = (\alpha,r_1,\ldots,r_\nu)$. For each $\gamma = (\eta,\phi,q)$ in Γ we shall denote the output to (2.1)-(2.3) at time $t \geq 0$ by $y = y(t;\gamma)$.

Identification of the system variables γ in (2.1)-(2.3) is based on input-output information. Given a piecewise continuous control input u defined on some time interval $[0,T]$, one samples the system at times $\{t_i\}$, $0 \le t_1 < t_2 < \ldots < t_M \le T$, to obtain observations $\{\hat{y}_i\}$, $\hat{y}_i \in R^k$, $i = 1,2,\ldots,M$. One can then perform a least squares fit to data (or seek a maximum likelihood estimator for γ). Formally, we may state this as follows:

Problem: Given the input u and observations $\{\hat{y}_i\}$ at times $\{t_i\}$, find $\gamma^* = (\eta^*,\phi^*,q^*)$ in Γ which minimizes the fit error

$$J(\gamma) = \frac{1}{2} \sum_{i=1}^{M} |y(t_i;\gamma) - \hat{y}_i|^2. \tag{2.5}$$

Remark 2: Whenever $r_\nu^* < r$, one only needs ϕ^* defined on $[-r^*,0]$ in order to obtain a solution to (2.1)-(2.3) (in practice, this is exactly what we shall obtain). However, we can view (η^*,ϕ^*) as an element of $\mathcal{S}$ by making a simple (arbitrary but definite) backward extension of ϕ^* to all of $[-r,0]$.

2.1 An abstract formulation of the I.D. problem

Let $r > 0$ be fixed and given as in the previous section and define $Z \equiv R^n \times L_2^n(-r,0)$. For $q = (\alpha,h) \in Q$ and $(\eta,\phi) \in Z$, define for $t \geq 0$ the mappings $S(t;q)\colon Z \to Z$ by

$$S(t;q)(\eta,\phi) = (x(t;\gamma),x_t(\gamma))$$

where x is the solution to (2.1) with $u \equiv 0$ and $x_t(\theta) = x(t+\theta)$, $-r \leq \theta \leq 0$. It is easily verified that for each q, $\{S(t;q)\}_{t\geq 0}$ is a strongly continuous semigroup of linear operators on Z. Furthermore, one finds [5] that the infinitesimal generator $\mathscr{A}(q)$, with domain

$$\mathscr{D}(\mathscr{A}(q)) = \mathscr{D} = \{(\eta,\phi) \in Z \mid \phi \in W_2^{(1)}(-r,0),\ \eta = \phi(0)\},$$

is given by

$$\mathscr{A}(q)(\phi(0),\phi) = (L(q)\phi,\dot{\phi}).$$

We note that for $q \in Q$, $\mathscr{D}(\mathscr{A}(q))$ does not depend on q itself. However, for $k > 1$, $\mathscr{D}(\mathscr{A}^k(q))$ <u>does</u> depend on q. For example, $\mathscr{D}(\mathscr{A}^2(q)) = \{(\phi(0),\phi) \mid \phi \in W_2^{(2)}(-r,0),\ \dot{\phi}(0) = L(q)\phi\}$.

If we define the operators $\hat{B}(\alpha)\colon R^m \to Z$ and $\hat{C}(\alpha)\colon Z \to R^k$ by $\hat{B}(\alpha)u = (B(\alpha)u,0)$ and $\hat{C}(\alpha)(\eta,\phi) = C(\alpha)\eta$, then the delay system (2.1)-(2.3) is formally equivalent to the abstract ordinary differential equation (ODE) system

$$(2.6) \qquad \dot{z}(t) = \mathscr{A}(q)z(t) + \hat{B}(\alpha)u(t), \ t \geq 0,$$

$$(2.7) \qquad z(0) = (\eta,\phi)$$

$$(2.8) \qquad y(t) = \hat{C}(\alpha)z(t) + D(\alpha)u(t).$$

As in the usual theory dealing with semigroups and abstract differential equations, a mild solution to (2.6)-(2.8) can be given by a variation of parameters formula. Specifically, (2.6)-(2.7) has the mild solution $z(t) = z(t;\gamma,u)$ given by

$$(2.9) \qquad z(t) = S(t;q)(\eta,\phi) + \int_0^t S(t-\sigma;q)\hat{B}(\alpha)u(\sigma)d\sigma.$$

It is a happy circumstance that (2.9) is actually equivalent to (2.1)-(2.2) in a strong sense as we now state precisely (for proof see [4], [5] or [6]).

Theorem 2.1. Let $x(\cdot;\gamma,u)$ denote the solution to (2.1)-(2.2) corresponding to $\gamma \in Z \times Q$ and $u \in L_{2,loc}$. Then, for all $t \geq 0$

$$z(t;\gamma,u) = (x(t;\gamma,u),x_t(\gamma,u)).$$

In view of the above equivalence results, the I.D. problem for (2.1)-(2.3) posed above can be reformulated in terms of an abstract I.D. problem. That is, given input u and observations $\{\hat{y}_i\}$ at times $\{t_i\}$, find $\gamma^* = (\eta^*,\phi^*,q^*)$ in Γ so as to minimize $J(\gamma)$ as given in (2.5) when now $y(t)$ is given by (2.8) and (2.9) in place of (2.3). Whether the problem is formulated in terms of (2.8), (2.9) or (2.1)-(2.3), it is clear that we are dealing with I.D. problems involving infinite dimensional state systems. Formulation in the framework of the Hilbert space Z only emphasizes this and is in no way an essential factor in the infinite dimensionality (and the associated difficulties) of the problem. Our main interests here are identification schemes that will result in computationally efficient algorithms. The approach we take is a classical one of the Ritz type. We shall choose a sequence of finite dimentional problems, each of which is defined on a finite dimensional state space X_N and approximates the original I.D. problem in Z. By appropriate choices of the sequence $\{X_N\}$ and the corresponding approximating problems, we hope to obtain a sequence of more easily solved problems with solutions $\gamma^N = (\eta^N, \phi^N, q^N)$ which converge to a solution γ^* of our original problem.

Fundamental to this endeavor is the convergence of the underlying approximating systems to the original system (2.9). Our formulation in a functional analytic framework will allow us to utilize abstract approximation theorems from semigroup theory (e.g., see [5]). The problems here, however, are a little different from the control problems of [5] where one chooses a sequence of subspaces $Z^N \subset Z$ on which to solve approximating control problems. The I.D. problems to be treated below pose some additional difficulties in that for each value of N, the "state" space changes. That is, the natural space for (2.9) with $q^N = (\alpha^N, r_1^N, \ldots, r_\nu^N)$ is $Z_N = R^n \times L_2^n(-r_\nu^N, 0)$ which, in addition to varying with N, is not a subspace of the original space

$Z = R^n \times L_2^n(-r,0)$. The approximating spaces X_N clearly should be chosen so that $X_N \subset Z_N$.

There are abstract approximation theorems (motivated by differencing schemes for partial differential equations and applications from probability theory) available in the literature in the case where $Z_N \not\subset Z$. For example, the original Lax, Trotter, Kato efforts [20], [26], [17] resulted in such theorems as did the later efforts of Kurtz [19]. However, all of these versions of the approximation results (and all others with which we are familiar) require the spaces X_N to approximate Z in the sense that there exist projection-like mappings $P_N: Z \to X_N$ which satisfy a norm convergence criterion $|P_N z|_{X_N} \to |z|_Z$ as $N \to \infty$. For the problems and approximations we shall discuss below such a criterion is not met (in general, one will <u>not</u> have $r_\nu^N \to r$, where r is the a priori chosen upper bound for the hereditary effects in the systems). We shall, therefore be obligated to state and prove an appropriate version of the abstract approximation results and this is done in the next section. The arguments used to establish this theorem are very similar to the standard ones found in the literature. One has a sequence of approximating infinitesimal generators (i.g.'s) A_N which converge in some sense to an i.g. A. This convergence is sufficient to imply convergence of the resolvents $R_\lambda(A_N)$ to $R_\lambda(A)$. These are the Laplace transforms of the corresponding semigroups $S_N(t), S(t)$ respectively and their convergence is enough to guarantee the desired convergence $S_N(t) \to S(t)$. We make this more precise in the next section.

3. An abstract approximation theorem

Let Z and Z_N, $N = 1,2,\ldots$, be Hilbert spaces with norms $|\cdot|$ and $|\cdot|_N$ respectively. Let X_N be a closed subspace of Z_N and $\pi_N: Z_N \to X_N$ be the canonical projection of Z_N onto X_N along $X_N^{\perp}$. Suppose $\mathscr{I}_N: Z \to Z_N$ is a mapping satisfying $\mathrm{Im}(\mathscr{I}_N) = Z_N$ and $|\mathscr{I}_N z|_N \leq |z|$ for $z \in Z$. Finally define $P_N: Z \to X_N$ by $P_N = \pi_N \mathscr{I}_N$. (In our discussions for the I.D. problem above, $Z = R^n \times L_2^n(-r,0)$, $Z_N = R^n \times L_2^n(-r_\nu^N,0)$, X_N is an approximating space such as the AVE spaces of [5] or the spline spaces of [10] - these will be discussed fully below. Finally, $\mathscr{I}_N$ is the operator that takes $z = (\eta,\phi)$ in Z into $\tilde{z} = (\eta,\tilde{\phi})$ where $\tilde{\phi}$ is the restriction of ϕ to $[-r_\nu^N,0]$. We note that in this case we would not expect to have $|P_N z|_N \to |z|$ for $z \in Z$ unless $r_\nu^N \to r$ and π_N itself has certain convergence properties.)

We adopt the following standard notation for the presentation of our fundamental approximation results. For a Hilbert space X, we write $B \in G(M,\beta)$ to mean $B: \mathscr{D}(B) \subset X \to X$ is the i.g. of a C_0-s.g. $\{T(t)\}$ satisfying $|T(t)| \leq Me^{\beta t}$. We also denote the resolvent $(\lambda I-B)^{-1}$ by $R_\lambda(B)$ and recall that $R_\lambda(B)x = \int_0^\infty e^{-\lambda\sigma}T(\sigma)x\,d\sigma$.

Theorem 3.1. Let Z, Z_N, X_N, and P_N be given as above. Suppose for some M,β we have $A_N \in G(M,\beta)$ on X_N and $A \in G(M,\beta)$ on Z. Further suppose there exists $\mathscr{D} \subset \mathscr{D}(A)$, $\mathscr{D}$ dense in Z such that

(3.1)
(i) $R_\lambda(A)\mathscr{D} \subset \mathscr{D}$ for $\mathrm{Re}\,\lambda > \beta$,
(ii) for every $z \in \mathscr{D}$, $|A_N P_N z - P_N A z|_N \to 0$ as $N \to \infty$.

Then for every $z \in Z$

$$(3.2) \qquad |S_N(t)P_N z - P_N S(t)z|_N \to 0 \quad \text{as} \quad N \to \infty,$$

and the convergence is uniform in t on compact intervals. Here A_N is the i.g. for $S_N(t)$, A the i.g. for $S(t)$.

Remark 3.1. Implicit in the statement and proof of Theorem 3.1 is the assumption that $P_N z \in \mathcal{D}(A_N)$ for every $z \in Z$. In our use of the theorem for I.D. schemes below, X_N will be finite dimensional and $\mathcal{D}(A_N) = X_N$. Indeed, we shall find A_N bounded with $S_N(t) = e^{A_N t}$.

Proof: Let λ be fixed throughout with $\text{Re}\,\lambda > \beta$ so that $R_\lambda(A_N)$, $R_\lambda(A)$ exist. We first establish that for every $y \in Z$

$$(3.3) \qquad |R_\lambda(A_N)P_N y - P_N R_\lambda(A)y|_N \leq \frac{M}{\text{Re}\,\lambda - \beta} |(A_N P_N - P_N A)R_\lambda(A)y|_N.$$

From the definition of the resolvent operator we have for any operator B

$$R_\lambda(B)B = BR_\lambda(B) = \lambda R_\lambda(B) - I.$$

In particular

$$R_\lambda(A_N)A_N P_N = \lambda R_\lambda(A_N)P_N - P_N$$
$$P_N A R_\lambda(A) = \lambda P_N R_\lambda(A) - P_N$$

so that

$$R_\lambda(A_N)A_NP_NR_\lambda(A) - R_\lambda(A_N)P_NAR_\lambda(A) = R_\lambda(A_N)P_N - P_NR_\lambda(A).$$

Hence for any $y \in Z$ we have

$$|R_\lambda(A_N)P_Ny - P_NR_\lambda(A)y|_N$$

$$= |R_\lambda(A_N)[A_NP_N - P_NA]R_\lambda(A)y|_N$$

$$\leq \frac{M}{\mathrm{Re}\ \lambda - \beta} |(A_NP_N - P_NA)R_\lambda(A)y|_N,$$

the last inequality following from the fact that $A_N \in G(M,\beta)$.

Next, for given $z \in \mathscr{D}$ where $\mathscr{D}$ is as in the hypotheses, define

$$F_N(\sigma) \equiv S_N(\sigma)P_Nz - P_NS(\sigma)z.$$

Then from (i), (ii) and (3.3) we conclude that for $\mathrm{Re}\ \lambda > \beta$,

$$|\mathscr{L}_\lambda[F_N]|_N \to 0 \quad \text{as} \quad N \to \infty$$

where $\mathscr{L}_\lambda$ is the Laplace transform. We observe that from the bounds on S_N, P_N, S the sequence $\{F_N\}$ is uniformly exponentially bounded, i.e.

$$|F_N(\sigma)| \leq 2Me^{\beta\sigma}|z|.$$

Finally, since

$$\frac{d}{d\tau} P_N S(\tau) z = P_N S(\tau) A z$$

and

$$\frac{d}{d\tau} S_N(\tau) P_N z = S_N(\tau) A_N P_N z,$$

a simple quadrature reveals

$$F_N(\sigma) = \int_0^\sigma [S_N(\tau) A_N P_N z - P_N S(\tau) A z] d\tau \tag{3.4}$$

and it follows that $\{F_N\}$ is a pointwise equicontinuous family on $[0,\infty)$. (From the convergence in (3.1)-(ii) and the bounds on S_N and S, one easily verifies that the integrand in (3.4) is uniformly exponentially bounded.) We are thus in a position to use a lemma due to Kurtz (Lemma 2.11, p. 359 of [19]) to conclude that $|F_N(\sigma)|_N \to 0$ as $N \to \infty$, uniformly on compact intervals. (Actually, the lemma as stated by Kurtz requires uniform boundedness of $\{F_N\}$ but a careful inspection of his proof will convince the reader that this requirement can be replaced by uniform exponential boundedness as we have here.)

We thus obtain the desired convergence (3.2) at least for each $z \in \mathscr{D}$. But then standard density arguments (the triangle inequality, bounds for S_N, P_N, and the density of $\mathscr{D}$ in Z) can be employed to establish the convergence for all $z \in Z$.

Remark 3.2. We note that in the above theorem we could have hypothesized $A_N \in G(M,\beta)$ on Z_N instead of on X_N without altering the proof. However, in the applications we have in mind we wish to obtain invariancy of $S_N(t) = e^{A_N t}$ on X_N (the space where our approximating systems will be defined and used). Thus, if we posit $A_N \in G(M,\beta)$ on Z_N we must add the additional hypothesis $\mathrm{Im}(A_N) \subset X_N \subset \mathcal{D}(A_N)$ in order to use the approximation result as we desire below.

Remark 3.3. One can clearly choose $X_N = Z_N$ (with π_N then the identity on Z_N or $Z_N = Z$ (and $\mathcal{I}_N$ the identity on Z) and obtain other versions of the approximation results. Again, our choice here is dictated by the application to be discussed below.

Remark 3.4. In the event one has $Z_N = X_N \subset Z$ and $P_N\colon Z \to Z_N$ satisfying $P_N z \to z$ for $z \in Z$, then the condition (3.1)-(ii) can be replaced by $|A_N P_N z - Az| \to 0$ and the conclusion (3.2) by $|S_N(t)P_N z - S(t)z| \to 0$. This then is essentially the version of the approximation theorem that we employed in previous efforts dealing with control problems [4], [5].

Corollary 3.1. Suppose the convergence in (3.1)-(ii) is $O(N^{-\delta})$ whenever z has the form $R_\lambda^2(A)y$, $S(t)R_\lambda(A)y$, and $S(t)R_\lambda^2(A)y$, $\lambda > \beta$, for a given $y \in Z$. Suppose further that the constants in $O(N^{-\delta})$ are uniform in t in the latter two cases. Then the convergence in (3.2) is also $O(N^{-\delta})$ whenever $z = R_\lambda^2(A)y$ for this y.

Proof: Using rather standard arguments (p. 87, [22]) one finds that

$$\frac{d}{d\sigma}\,[S_N(t-\sigma)R_\lambda(A_N)P_NS(\sigma)R_\lambda(A)x]$$

$$= S_N(t-\sigma)[P_NR_\lambda(A) - R_\lambda(A_N)P_N]S(\sigma)x$$

for arbitrary $x \in Z$ and $\lambda > \beta$. Hence, we have

$$(3.5)\quad \int_0^t S_N(t-\sigma)[P_NR_\lambda(A) - R_\lambda(A_N)P_N]S(\sigma)x\,d\sigma$$

$$= R_\lambda(A_N)[P_NS(t) - S_N(t)P_N]R_\lambda(A)x.$$

Using (3.5) and (3.3) we have for any $y \in Z$

$$|[S_N(t)P_N - P_NS(t)]R_\lambda^2(A)y|_N$$

$$\leq |S_N(t)[P_NR_\lambda(A) - R_\lambda(A_N)P_N]R_\lambda(A)y|_N$$

$$+ |R_\lambda(A_N)[S_N(t)P_N - P_NS(t)]R_\lambda(A)y|_N$$

$$+ |[R_\lambda(A_N)P_N - P_NR_\lambda(A)]S(t)R_\lambda(A)y|_N$$

$$\leq Me^{\beta t}|[P_NR_\lambda(A) - R_\lambda(A_N)P_N]R_\lambda(A)y|_N$$

$$+ \int_0^t |S_N(t-\sigma)||\{P_NR_\lambda(A) - R_\lambda(A_N)P_N\}S(\sigma)y|_N d\sigma$$

$$+ |\{R_\lambda(A_N)P_N - P_NR_\lambda(A)\}S(t)R_\lambda(A)y|_N$$

$$\leq Me^{\beta t} \frac{M}{\lambda-\beta} |(P_N A - A_N P_N) R_\lambda^2(A) y|_N$$

$$+ \int_0^t Me^{\beta(t-\sigma)} \frac{M}{\lambda-\beta} |(P_N A - A_N P_N) S(\sigma) R_\lambda(A) y|_N \, d\sigma$$

$$+ \frac{M}{\lambda-\beta} |(P_N A - A_N P_N) S(t) R_\lambda^2(A) y|_N .$$

Thus, if y is chosen as in the hypothesis of the corollary, the conclusion follows immediately.

Theorem 3.2. Let $\mathscr{B} \subset \mathscr{D}(A^2)$ satisfy the following:

(i) For each $z \in \mathscr{B}$ there exists $k = k(z)$ such that

$$|(A_N P_N - P_N A) z|_N \leq k/N^\delta, \quad N = 1,2,\ldots .$$

(ii) There exists $\mathscr{B}_1 \subset \mathscr{B}$ such that $z \in \mathscr{B}_1$ implies

(a) $S(t)z \in \mathscr{B}$, $0 \leq t \leq T$,

(b) $S(t)(\lambda I - A)z \in \mathscr{B}$, $\lambda > \beta$, $0 \leq t \leq T$

and furthermore the constants guaranteed by (i) for (a), (b) can be chosen independent of t.

Then for $z \in \mathscr{B}_1$ we have there exists $k(z)$ such that

$$|[S_N(t) P_N - P_N S(t)] z|_N \leq \tilde{k}(z)/N^\delta$$

for $0 \leq t \leq T$.

Proof: Let $z \in \mathscr{B}_1$ and define $y = (\lambda I - A)^2 z$. Then $R_\lambda^2(A)y = z$. Furthermore, by (ii) we have $S(t)R^2(A)y \in \mathscr{B}$ and $S(t)R_\lambda(A)y \in \mathscr{B}$ with the constants in the $O(N^{-\delta})$ estimates uniform in t. It follows that the hypotheses of Corollary 3.1 are satisfied and hence we reach the desired conclusion since $z = R_\lambda^2(A)y$.

4. Identification schemes for delay systems

Let π^0,π^1 be the coordinate projections of $Z = R^n \times L_2^n(-r,0)$ onto R^n, $L_2^n(-r,0)$ respectively. We recall that the I.D. problem of §2 can be written:

($\mathscr{P}$): Given input u and observations $\{\hat{y}_i\}$ at $\{t_i\}_{i=1}^M$, find $\gamma^* = (\eta^*,\phi^*,q^*)$ in Γ so as to minimize

$$J(\gamma) = \frac{1}{2}\sum_{i=1}^{M} |y(t_i;\gamma) - \hat{y}_i|^2$$

where y is the solution of (2.8), (2.9). That is,

$$\begin{aligned} y(t;\gamma) &= \hat{C}(\alpha)z(t;\gamma,u) + D(\alpha)u(t) \\ &= C(\alpha)\pi^0 z(t;\gamma,u) + D(\alpha)u(t) \\ &= C(\alpha)x(t;\gamma,u) + D(\alpha)u(t) \end{aligned}$$

Thus the identification problem can be viewed in a state-space Z, parameter space Γ setting. This will lead to a sequence of approximate I.D. problems if we approximate Z by a sequence of spaces X_N.

We henceforth assume that we have in hand a sequence $\{r_\nu^N\}$, $0 \le r_\nu^N \le r$, and take closed $X_N \subset Z_N \equiv R^n \times L_2^n(-r^N,0)$. Then we define the operator $\mathscr{F}_N: Z \to Z_N$ as the operator that truncates $\pi^1 z$ to the interval $[-r_\nu^N,0]$ and then denote by $\mathscr{F}_N^\dagger$ the Moore-Penrose [21] pseudo inverse $\mathscr{F}_N^\dagger: Z_N \to Z$. In this case if $z = (\eta,\psi) \in Z_N$, then $\mathscr{F}_N^\dagger(\eta,\psi) = (\eta,\phi)$ where $\phi = \psi$ on $[-r_\nu^N,0]$, $\phi = 0$ on $[-r,-r_\nu^N)$.

Let $\{S_N(t)\}$ be a C_0-semigroup on X_N with i.g. $A_N \in G(M,\beta)$ and let $P_N\colon Z \to X_N$ be as given in §3. We define $\Gamma_N \subset X_N \times Q$ by $\Gamma_N = (P_N\mathcal{S}) \times Q$ where $\mathcal{S}$ is the given initial data set. We make the following standing assumption on $\mathcal{S}$ and Q.

<u>Assumption 4.1</u>. Q and $\mathcal{S}$ are compact and $\mathcal{I}_N^{\dagger}\mathcal{I}_N(\mathcal{S}) \subset \mathcal{S}$ for all N.

From this assumption and the continuity of P_N, we readily obtain that Γ_N is compact. Given $\gamma^N = (z_0^N, q^N) \in \Gamma_N$ we define

$$(4.1) \qquad z^N(t;\gamma^N,u) \equiv S_N(t)z_0^N + \int_0^t S_N(t-\sigma)P_N\hat{B}(\alpha^N)u(\sigma)d\sigma$$

and

$$(4.2) \qquad y^N(t;\gamma^N) \equiv \hat{C}(\alpha^N)z^N(t;\gamma^N,u) + D(\alpha^N)u(t)$$
$$= C(\alpha^N)\pi^0 z^N(t;\gamma^N,u) + D(\alpha^N)u(t).$$

The approximate I.D. problems are:

$(\mathcal{P}_N)$: Given input u and observations $\{\hat{y}_i\}$, find $\bar{\gamma}^N \in \Gamma_N$ so as to minimize

$$J^N(\gamma^N) \equiv \sum_{i=1}^{M} |y^N(t_i;\gamma^N) - \hat{y}_i|^2.$$

We note that solutions to $(\mathcal{P}_N)$ exist since the mapping $\xi \to J^N(\xi)$ is continuous on Γ_N and Γ_N is compact.

<u>Theorem 4.1</u>. Suppose $\overline{\gamma}^N = (\overline{z}_0^N, \overline{q}^N)$ is a sequence of solutions to the problems $(\mathscr{P}_N)$ and that there exists $\overline{\gamma} \in \Gamma$ such that $\overline{\gamma}^N \to \overline{\gamma}$ in the sense (a) $\overline{q}^N \to \overline{q}$ in $R^{\mu+\nu}$, (b) $\mathscr{I}_N^\dagger \overline{z}_0^N \to \overline{z}_0$ in Z. Suppose further that $A_N = A_N(\overline{q}^N)$, $A = A(\overline{q})$ satisfy the conditions and hypotheses of Theorem 3.1. Then

$$|P_N z(t;\overline{\gamma},u) - z^N(t;\overline{\gamma}^N,u)|_N \to 0 \quad \text{as} \quad N \to \infty,$$

uniformly in t on compact intervals, where

$$(4.3) \quad z(t;\overline{\gamma},u) \equiv S(t)\overline{z}_0 + \int_0^t S(t-\sigma)\hat{B}(\overline{\alpha})u(\sigma)d\sigma.$$

<u>Proof</u>: From the hypotheses and Theorem 3.1 we have immediately that $|S_N(t)P_N z - P_N S(t)z|_N \to 0$, uniformly on compact intervals, for all $z \in Z$. Therefore

$$|S_N(t)\overline{z}_0^N - P_N S(t)\overline{z}_0|_N = |S_N(t)P_N \mathscr{I}_N^\dagger \overline{z}_0^N - P_N S(t)\overline{z}_0|_N$$

$$\leq |S_N(t)[P_N \mathscr{I}_N^\dagger \overline{z}_0^N - P_N\overline{z}_0]|_N + |S_N(t)P_N\overline{z}_0 - P_N S(t)\overline{z}_0|_N$$

$$\leq Me^{\beta t}|\mathscr{I}_N^\dagger \overline{z}_0^N - \overline{z}_0| + |S_N(t)P_N\overline{z}_0 - P_N S(t)\overline{z}_0|_N.$$

The first term approaches 0 by (b) as does the second from our preceding remark. Next, consider

$$\left|\int_0^t S_N(t-\sigma)P_N\hat{B}(\bar{\alpha}^N)u(\sigma)d\sigma - P_N\int_0^t S(t-\sigma)\hat{B}(\bar{\alpha})u(\sigma)d\sigma\right|_N$$

$$= \left|\int_0^t [S_N(t-\sigma)P_N\hat{B}(\bar{\alpha}^N)u(\sigma) - P_N S(t-\sigma)\hat{B}(\bar{\alpha})u(\sigma)]d\sigma\right|_N$$

$$\leq \int_0^t |S_N(t-\sigma)P_N[\hat{B}(\bar{\alpha}^N) - \hat{B}(\bar{\alpha})]u(\sigma)|_N d\sigma$$

$$+ \int_0^t |[S_N(t-\sigma)P_N - P_N S(t-\sigma)]\hat{B}(\bar{\alpha})u(\sigma)|_N d\sigma.$$

The results of Theorem 3.1, the continuity of $\hat{B}$, and dominated convergence yield convergence of these terms to 0, uniformly in t. The desired conclusion follows immediately from these estimates.

<u>Corollary 4.1</u>. *Suppose* $P_N: Z \to X_N$ satisfies

(4.4) $\qquad \pi^0(P_N z) \to \pi^0 z$ in R^n for each $z \in Z$.

Then under the assumptions of Theorem 4.1 we have

$$y^N(t;\bar{\gamma}^N) \to y(t;\bar{\gamma}) \quad \text{for each } t.$$

<u>Proof</u>: Recall

$$y(t;\bar{\gamma}) = \hat{C}(\bar{\alpha})z(t;\bar{\gamma},u) + D(\bar{\alpha})u(t)$$

$$= C(\bar{\alpha})\pi^0 z(t;\bar{\gamma},u) + D(\bar{\alpha})u(t)$$

while

$$y^N(t;\bar{\gamma}^N) = C(\bar{\alpha}^N)\pi^0 z^N(t;\bar{\gamma}^N,u) + D(\bar{\alpha}^N)u(t).$$

The claimed result follows at once from the result of Theorem 4.1

$$|\pi^0 z^N(t;\bar{\gamma}^N,u) - \pi^0 P_N z(t;\bar{\gamma},u)|_{R^n} \to 0,$$

and (4.4) which yields

$$|\pi^0 P_N z(t;\bar{\gamma},u) - \pi^0 z(t;\bar{\gamma},u)|_{R^n} \to 0.$$

We observe that in (4.1) we define z^N for initial data in X_N. However, one can define an analogue for initial data given in $\mathscr{S}$. In particular for fixed $\gamma = (z_0,q) \in \Gamma = \mathscr{S} \times Q$ we define

$$(4.5) \quad \tilde{z}^N(t;\gamma,u) \equiv S_N(t)P_N z_0 + \int_0^t S_N(t-\sigma)P_N\hat{B}(\alpha)u(\sigma)d\sigma$$

where $A_N = A_N(q)$, $A = A(q)$ are i.g.'s for S_N,S. If one then assumes that $A_N(q)$,$A(q)$ satisfy the hypotheses of Theorem 3.1 so that $|P_N S(t)z - S_N(t)P_N z|_N \to 0$, one can prove in almost exactly the same manner as that for Theorem 4.1 above that

$$|\tilde{z}^N(t;\gamma,u) - P_N z(t;\gamma,u)|_N \to 0$$

for $\gamma \in \Gamma$. Defining $\tilde{y}^N(t;\gamma)$ as in (4.2) except with $\tilde{z}^N(t;\gamma,u)$ of (4.5) in place of $z^N(t;\gamma^N,u)$, we have under hypothesis (4.4) the analogue of the results of Corollary 4.1:

(4.6) $$\tilde{y}^N(t;\gamma) \to y(t;\gamma)$$

for each fixed $\gamma \in \Gamma$.

Theorem 4.2. Suppose $\{\bar{\gamma}^N\}$ is a sequence of solutions of the approximate problems $(\mathcal{P}_N)$ under Assumption 4.1. Then there exist $\bar{\gamma} \in \Gamma$ and a subsequence $\{\bar{\gamma}^{N_k}\}$ such that $\bar{\gamma}^{N_k} \to \bar{\gamma}$ in the sense of Theorem 4.1(a), (b). If $A_N(\bar{q}^N), A(\bar{q})$ satisfy the hypotheses of Theorem 3.1, then $\bar{\gamma}$ is a solution for the problem $(\mathcal{P})$.

Proof: First we have $\bar{\gamma}^N = (\bar{z}_0^N, \bar{q}^N) \in (P_N\mathcal{S}) \times Q$. Defining $\bar{z}^N \equiv \mathcal{I}_N^+ \bar{z}_0^N$, we have $\bar{z}^N \in \mathcal{S}$, $\mathcal{S}$ compact, so that there exists a convergent subsequence, say $\{\bar{z}^{N_k}\}$, converging to some $\bar{z}_0$ in $\mathcal{S}$; i.e., $\mathcal{I}_{N_k}^+ \bar{z}_0^{N_k} \to \bar{z}_0$ in $\mathcal{S}$. From the compactness of Q, we have that $\{\bar{q}^{N_k}\}$ possesses a convergent subsequence with $\bar{q}^{N_{k_j}} \to \bar{q}$ for some $\bar{q} \in Q$. Defining $\bar{\gamma} = (\bar{z}_0, \bar{q}) \in \Gamma = \mathcal{S} \times Q$, and reindexing we thus have a sequence $\{\bar{\gamma}^{N_j}\}$ that converges in the sense of Theorem 4.1(a),(b) to $\bar{\gamma}$. Furthermore, it follows from Theorem 4.1, Corollary 4.1 and the remarks involving (4.5) and (4.6) that for any $\gamma = (z_0, q) \in \Gamma$ we have

$$J(\bar{\gamma}) = \lim_{N_j \to \infty} J^{N_j}(\bar{\gamma}^{N_j}) \quad \text{(Corollary 4.1 yields } y^{N_j}(t;\bar{\gamma}^{N_j}) \to y(t;\bar{\gamma})\text{)}.$$

But we find

$$\lim_{N_j \to \infty} J^{N_j}(\bar{\gamma}^{N_j}) \leq \lim_{N_j \to \infty} J^{N_j}((P_{N_j} z_0, q))$$

$$= \lim \left[\sum_{i=1}^{M} |y^{N_j}(t_i;(P_{N_j} z_0,q)) - \hat{y}_i|^2 \right].$$

But $y^{N_j}(t_i;(P_{N_j} z_0,q))$ given by (4.1) is exactly the same as $\tilde{y}^{N_j}(t_i;\gamma)$, $\gamma = (z_0,q)$, where $\tilde{y}^N$ is defined as in (4.5), (4.6) and hence the latter term is the same as

$$\lim_{N_j \to \infty} \left[\sum_{i=1}^{M} |\tilde{y}^{N_j}(t_i;\gamma) - \hat{y}_i|^2 \right] = J(\gamma).$$

Thus, $\bar{\gamma}$ is a solution for $(\mathscr{P})$.

We turn next to a discussion of particular schemes which fit into the theoretical framework developed above. Throughout our presentation we shall assume that we are given a sequence $\gamma^N = (\eta^N, \phi^N, q^N)$ when $q^N = (\alpha^N, h^N) = (\alpha^N, r_1^N, \ldots, r_\nu^N) \in Q$, with $0 < r_1^N < r_2^N < \ldots < r_\nu^N \leq r$, and $q^N \to \bar{q} = (\bar{\alpha}, \bar{h}) = (\bar{\alpha}, \bar{r}_1, \ldots, \bar{r}_\nu) \in Q$. We recall that for the systems under discussion we have the operator $\mathscr{A} = \mathscr{A}(\bar{q})$ defined on $\mathscr{D} = \{(\phi(0),\phi) \mid \phi \in W_2^{(1)}(-r,0)\}$ given by

$$\mathscr{A}(\bar{q})(\phi(0),\phi) = (L(\bar{q})\phi, D\phi)$$

where the operator L is defined in (2.4). Hereafter we shall use the notation $D\phi$ in place of $\dot{\phi}$ in context where confusion might arise otherwise.

We summarize for future reference the conditions that our approximating schemes must satisfy:

(4.7) X_N is a closed subspace of $Z_N = R^n \times L_2^n(-r_\nu^N, 0)$, π_N is the canonical projection of Z_N onto X_N, $P_N = \pi_N \mathscr{I}_N$ and $\pi^0(P_N z) \to \pi^0 z$ for all $z \in Z$.

(4.8) There exist constants M and β such that $\mathscr{A}_N = \mathscr{A}_N(q^N)$ and $\mathscr{A} = \mathscr{A}(\bar{q})$ are in $G(M,\beta)$ on X_N and Z respectively.

(4.9) There exists $\mathscr{D}_1 \subset \mathscr{D} = \mathscr{D}(\mathscr{A}(\bar{q}))$, $\mathscr{D}_1$ dense in Z, such that

(i) $R_\lambda(\mathscr{A}(\bar{q}))\, \mathscr{D}_1 \subset \mathscr{D}_1$ for $\lambda > \beta$

(ii) $|\mathscr{A}_N P_N z - P_N \mathscr{A} z|_N \to 0$ as $N \to \infty$ for $z \in \mathscr{D}_1$.

In our discussions below we shall refer to (4.8) as the stability condition while (4.9) will be called the consistency condition. Our first scheme will be based on the averaging approximation developed in some detail in [5] while the second scheme utilizes spline approximations as formulated in [10].

5. The averaging approximation scheme

This identification scheme is defined using the "averaging" type approximations as discussed in [4], [5] and many of the arguments to verify that conditions (4.7), (4.8), (4.9) are satisfied are only slight modifications of those found in [5]. Given $q^N = (\alpha^N, r_1^N, \ldots, r_\nu^N)$, we partition $[-r_\nu^N, 0]$ into subintervals $[t_j^N, t_{j-1}^N]$ where $t_j^N \equiv -jr_\nu^N/N$, $j = 0,1,\ldots,N$. Let χ_j^N denote the characteristic function of $[t_j^N, t_{j-1}^N)$ for $j = 2,3,\ldots,N$, with χ_1^N the characteristic function for $[t_1^N, t_0^N] = [-r_\nu^N/N, 0]$. Define for $(\eta,\phi) \in Z$

$$\phi_j^N \equiv \frac{N}{r_\nu^N} \int_{t_j^N}^{t_{j-1}^N} \phi(s)ds, \qquad j = 1,2,\ldots,N, \tag{5.1}$$

$$\phi_0^N \equiv \eta.$$

We define the closed subspaces X_N of Z_N by

$$X_N \equiv \{(\eta,\psi) \in Z_N \mid \eta \in R^n, \quad \psi = \sum_{j=1}^{N} v_j^N \chi_j^N, v_j^N \in R^n\}.$$

The projection π_N of Z_N onto X_N is then given by

$$\pi_N(\eta,\phi) = (\eta, \sum_{j=1}^{N} \phi_j^N \chi_j^N). \tag{5.2}$$

With these definitions it is immediately obvious that (4.7) is satisfied.

For the operator L given by (2.4), we define the approximating operator $L_N(q^N)\colon X_N \to R^n$ by

$$(5.3)\quad L_N(q^N)(\eta, \sum_1^N v_j \chi_j^N) \equiv A_0(\alpha^N)\eta + \sum_{i=1}^{\nu} \sum_{j=1}^{N} A_i(\alpha^N) v_j \chi_j^N(-r_i^N) + \sum_{j=1}^{N} K_j^N(\alpha^N) v_j$$

where

$$(5.4)\quad K_j^N(\alpha) \equiv \int_{t_j^N}^{t_{j-1}^N} K(\alpha,\theta)\,d\theta .$$

Next, we define $D_N\colon X_N \to L_2^n(-r_\nu^N,0)$ by

$$(5.5)\quad D_N(\eta, \sum_1^N v_j \chi_j^N) \equiv \sum_{j=1}^{N} \frac{N}{r_\nu^N} \{v_{j-1} - v_j\} \chi_j^N$$

where $v_0 \equiv \eta$. Finally, we define $\mathscr{A}_N(q^N)\colon X_N \to X_N$ by

$$(5.6)\quad \mathscr{A}_N(q^N)(\eta,\psi) \equiv (L_N(q^N)(\eta,\psi),\ D_N(\eta,\psi)).$$

The proof that $\mathscr{A}_N(q^N) \in G(M,\beta)$ on X_N for some M and β independent of N is essentially given in [5] (see p. 183 and p. 186). One first argues that there is an equivalent inner product $\langle\ ,\ \rangle_{g^N}$ on X_N such that $\langle \mathscr{A}_N(q^N)z,z\rangle_{g^N} \le \beta(q^N)\langle z,z\rangle_{g^N}$ for all $z \in X_N$. As in [5], we define, for given $r_1^N, r_2^N, \ldots, r_\nu^N$, the index set $J^N = \{j_1^N,\ldots,j_{\nu-1}^N\}$ where j_i^N is the index such that $-r_i^N \in [t_{j_i^N}^N, t_{j_i^N-1}^N)$, $i = 1,2,\ldots,\nu-1$, and $j_\nu^N \equiv N$. We next define numbers a_j^N by $a_N^N = 1$ and, for $j = N-1, N-2, \ldots, 1$,

$$a_j^N = \begin{cases} a_{j+1}^N + 1 & \text{if } j \in J^N \\ a_{j+1}^N & \text{if } j \notin J^N. \end{cases}$$

Then define the nondecreasing piecewise constant weighting function g^N by $g^N(\theta) = a_j^N$, $t_j^N \le \theta < t_{j-1}^N$, $j = 1,2,\ldots,N$. Finally, we take $Z_N(g^N)$ and $X_N(g^N)$ as the spaces Z_N, X_N respectively with equivalent topology generated by the inner product

$$\text{(5.7)} \qquad \langle (\eta,\phi)\ ,(\zeta,\psi)\rangle_{g^N} \equiv \langle \eta,\zeta\rangle_{R^n} + \int_{-r_\nu^N}^{0} \phi\psi g^N .$$

If we then consider $(\eta,\psi) = (\eta, \sum_{j=1}^{N} v_j\chi_j^N) \in X_N$ (and define $v_0 \equiv \eta$), we find in a straightforward manner using estimates similar to those in [5, p. 186] that

$$\text{(5.8)} \quad \langle \mathscr{A}_N(q^N)(\eta,\psi),(\eta,\psi)\rangle_{g^N} \le \{|A_0(\alpha^n)| + \frac{1}{2}\sum_{i=1}^{\nu} |A_i(\alpha^N)|^2\}|\eta|^2$$
$$+ \frac{1}{2}\sum_{i=1}^{\nu} |v_{j_i}|^2 + \sum_{j=1}^{N} |K_j^N(\alpha^N)||v_j||\eta| + \sum_{j=1}^{N} \langle v_{j-1}-v_j, v_j\rangle\, a_j^N .$$

Noting that for $\psi = \Sigma\, v_j\chi_j^N$,

$$|(\eta,\psi)|_N^2 = |\eta|^2 + \sum_{j=1}^{N} \frac{r_\nu^N}{N} |v_j|^2 ,$$

we find

$$\sum_{j=1}^{N} |K_j^N(\alpha^N)|\ |v_j|\ |\eta| = \sum_{j=1}^{N} \left| \left(\frac{N}{r_\nu^N}\right)^{1/2} \int_{t_j^N}^{t_{j-1}^N} K(\alpha^N,\theta)\,d\theta \right| |\eta| \left| \left(\frac{r_\nu^N}{N}\right)^{1/2} v_j \right|$$

$$\le \sum_{j=1}^{N} \left\{ \frac{1}{2}\ |\eta|^2 \left| \left(\frac{N}{r_\nu^N}\right)^{1/2} \int_{t_j^N}^{t_{j-1}^N} K(\alpha^N,\theta)\,d\theta \right|^2 + \frac{1}{2}\ \frac{r_\nu^N}{N}\ |v_j|^2 \right\}$$

$$\le \sum_{j=1}^{N} \left\{ \frac{1}{2}\ |\eta|^2 \left(\frac{N}{r_\nu^N} \right) (t_{j-1}^N - t_j^N) \int_{t_j^N}^{t_{j-1}^N} |K(\alpha^N,\theta)|^2 d\theta + \frac{1}{2}\ \frac{r_\nu^N}{N}\ |v_j|^2 \right\}$$

$$\le \frac{1}{2}\ |\eta|^2 \int_{-r_\nu^N}^{0} |K(\alpha^N,\theta)|^2 d\theta + \frac{1}{2} \sum_{j=1}^{N} \frac{r_\nu^N}{N}\ |v_j|^2$$

$$\le \left[\frac{1}{2} + \frac{1}{2} \int_{-r_\nu^N}^{0} |K(\alpha^N,\theta)|^2 d\theta \right] |(\eta,\psi)|_N^2 .$$

Next observe that

$$\sum_{j=1}^{N} \langle v_{j-1} - v_j, v_j \rangle a_j^N \le \sum_{j=1}^{N} \{ \tfrac{1}{2}\ |v_{j-1}|^2 + \tfrac{1}{2}\ |v_j|^2 - |v_j|^2 \} a_j^N$$

$$= \sum_{j=1}^{N} \{ \tfrac{1}{2} |v_{j-1}|^2 - \tfrac{1}{2} |v_j|^2 \} a_j^N$$

$$= \frac{1}{2}\ |\eta|^2 a_1^N + \frac{1}{2} \sum_{j=1}^{N-1} (a_{j+1}^N - a_j^N) |v_j|^2 - \frac{1}{2}\ |v_N|^2 a_N^N$$

$$= \frac{1}{2}\ \nu |\eta|^2 - \frac{1}{2} \sum_{i=1}^{\nu} |v_{j_i}|^2 .$$

Using these estimates in (5.8) we finally obtain

$$\langle \mathscr{A}_N(q^N)(\eta,\psi),(\eta,\psi)\rangle_{g^N} \le \beta(q^N)|(\eta,\psi)|_N^2$$

where

$$\beta(q^N) \equiv |A_0(\alpha^N)| + \frac{1}{2}\sum_{i=1}^{\nu}|A_i(\alpha^N)|^2 + \frac{1}{2}\int_{-r_\nu^N}^{0}|K(\alpha^N,\theta)|^2 d\theta + \frac{\nu+1}{2} .$$

From the continuity assumptions made in §2 (see (2.4)) and the fact that $q^N \in Q$, Q compact, we have existence of $\tilde{\beta}$ such that $\beta(q^N) \le \tilde{\beta}$ for all N. Finally, since the X_N and $X_N(g^N)$ norms are equivalent independent of N, one finds (again see p. 186 of [5]) $\mathscr{A}_N(q^N) \in G(M,\tilde{\beta})$ on X_N. Since $\mathscr{A}(\bar{q})$ is the i.g. for a C_0-semigroup it also satisfies the requirement $\mathscr{A}(\bar{q}) \in G(M_1,\beta_1)$ on Z for some M_1 and β_1. It follows that the stability condition (4.8) is satisfied for our averaging approximations.

We next consider the consistency criteria (4.9). We take $\mathscr{D}_1 \equiv \{(\phi(0),\phi)|\phi$ is C^1 on $[-r,0]\}$. Then clearly $\mathscr{D}_1$ is dense in Z and $\mathscr{D}_1 \subset \mathscr{D}(\mathscr{A}(\bar{q}))$. Furthermore, $R_\lambda(A(\bar{q}))\mathscr{D}_1 \subset \mathscr{D}(\mathscr{A}^2(\bar{q})) \subset \mathscr{D}_1$ (see §2 above) so that (4.9)-(i) is satisfied for this choice of $\mathscr{D}_1$.

It remains to establish (4.9)-(ii). Given $z = (\phi(0),\phi)$ in $\mathscr{D}_1$ we observe that

$$P_N\mathscr{A}(\bar{q})z = (L(\bar{q})\phi, \sum_{j=1}^{N}(D\phi)_j^N\chi_j^N) \tag{5.9}$$

where

$$(D\phi)_j^N \equiv \frac{N}{r_\nu^N} \int_{t_j^N}^{t_{j-1}^N} D\phi(s)ds = \frac{N}{r_\nu^N} [\phi(t_{j-1}^N) - \phi(t_j^N)],$$

while

$$(5.10) \quad \mathscr{A}_N(q^N)P_N z = \mathscr{A}_N(q^N)(\phi_0^N, \sum_{j=1}^N \phi_j^N \chi_j^N) = (L_N(q^N)P_N z, D_N P_N z)$$

where $L_N(q^N)$ and D_N are given by (5.3) and (5.5). In view of (5.9), (5.10) and (1.4), it thus suffices to show

$$(5.11) \quad A_0(\alpha^N)\phi(0) + \sum_{i=1}^{\nu} \sum_{j=1}^{N} A_i(\alpha^N)\phi_j^N \chi_j^N(-r_i^N) + \sum_{j=1}^{N} K_j^N(\alpha^N)\phi_j^N$$

$$\xrightarrow{R^n} A_0(\overline{\alpha})\phi(0) + \sum_{i=1}^{\nu} A_i(\overline{\alpha})\phi(-\overline{r}_i) + \int_{-\overline{r}_\nu}^{0} K(\overline{\alpha},\theta)\phi(\theta)d\theta$$

and

$$(5.12) \quad \int_{-r_\nu^N}^{0} \left| \sum_{j=1}^{N} [\frac{N}{r_\nu^N} (\phi_{j-1}^N - \phi_j^N) - (D\phi)_j^N]\chi_j^N \right|^2 \to 0$$

as $N \to \infty$ (and $r_i^N \to \overline{r}_i$).

Consider (5.12) first and write this integral as

$$\int_{\frac{-r_\nu^N}{N}}^{0} \left| \frac{N}{r_\nu^N} (\phi_0^N - \phi_1^N - [\phi(0) - \phi(\frac{-r_\nu^N}{N})]) \right|^2$$

$$+ \sum_{j=2}^{N} \int_{t_j^N}^{t_{j-1}^N} \left| \frac{N}{r_\nu^N} (\phi_{j-1}^N - \phi_j^N - [\phi(t_{j-1}^N) - (t_j^N)]) \right|^2$$

$$= T_1^N + T_2^N.$$

Using an analogue of (3.18), p. 177 of [5] with r replaced by r_ν^N, estimates similar to those of [5] yield

$$T_2^N \leq r_\nu^N \left| \sup_{1 \leq j \leq N} 2\mathscr{S}_j^N \right|^2$$

where, as in [5], we define

$$\mathscr{S}_j^N \equiv \sup\{|\dot{\phi}(\theta)-\dot{\phi}(s)| \,\big|\, s,\ \theta \in [t_j^N, t_{j-1}^N]\}.$$

Use of the analogue of (3.18) of [5] in T_1^N allows us to write (after arguing in much the same manner as done on p. 177 of [5])

$$T_1^N \leq \frac{r_\nu^N}{N} \left\{\frac{1}{2} \, |\dot{\phi}(\frac{-r_\nu^N}{N})| + \frac{1}{2}\, \mathscr{S}_1^N\right\}^2 .$$

Since $\mathscr{S}_j^N \to 0$ as $N \to \infty$, uniformly in j, we conclude that (5.11) obtains.

We remark that if $\dot{\phi}(0) = 0$ and $\phi \in W_\infty^{(2)}(-r,0)$, then

$$\left|\dot{\phi}(\frac{-r_\nu^N}{N})\right| = \left|\dot{\phi}(\frac{-r_\nu^N}{N}) - \dot{\phi}(0)\right| \leq \sup|\ddot{\phi}(\theta)|\frac{r_\nu^N}{N} ,$$

and, since $\mathscr{S}_j^N$ is $O(\frac{r_\nu^N}{N})$ - see p. 178 of [5], we find that the convergence in (5.12) is $O(\frac{1}{N^2})$ or that the convergence in the second component (L_2 component) of (4.9)-(ii) is $O(\frac{1}{N})$.

Returning to (5.11) and recalling that $A_i(\alpha^N) \to A_i(\bar{\alpha})$, we see that to establish (5.11), we only need show

$$(5.13)\qquad \sum_{j=1}^{N} \phi_j^N \chi_j^N(-r_i^N) \to \phi(-\bar{r}_i), \quad i = 1,2,\ldots,\nu,$$

and

$$(5.14)\qquad \sum_{j=1}^{N} K_j^N(\alpha^N)\phi_j^N \to \int_{-\bar{r}_\nu}^{0} K(\bar{\alpha},\sigma)\phi(\sigma)d\sigma.$$

For ϕ in C^1 on $[-r,0]$ we have, for $\theta \in [t_j^N, t_{j-1}^N)$,

$$(5.15)\qquad |\phi_j^N - \phi(\theta)| = \left|\frac{N}{r_\nu^N}\int_{t_j^N}^{t_{j-1}^N}[\phi(\sigma)-\phi(\theta)]d\sigma\right| \le \frac{N}{r_\nu^N}\int_{t_j^N}^{t_{j-1}^N}|\dot{\phi}|_\infty|\sigma-\theta|d\sigma$$

$$\le \frac{N}{r_\nu^N}|\dot{\phi}|_\infty\left(\frac{r_\nu^N}{N}\right)^2 = |\dot{\phi}|_\infty \frac{r_\nu^N}{N}.$$

From this, it follows immediately that

$$(5.16)\qquad |\Sigma\, \phi_j^N \chi_j^N - \phi|^2_{L_2(-r_\nu^N,0)} \le (|\dot{\phi}|_\infty)^2\left(\frac{r_\nu^N}{N}\right)^2.$$

For $j_i = j_i^N$ chosen so that $-r_i^N \in [t_{j_i}^N, t_{j_i-1}^N)$ we find using (5.15)

$$|\Sigma\, \phi_j^N \chi_j^N(-r_i^N) - \phi(-\bar{r}_i)| \le |\phi_{j_i}^N - \phi(-r_i^N)| + |\phi(-r_i^N) - \phi(-\bar{r}_i)|$$
$$\le |\dot{\phi}|_\infty \frac{r_\nu^N}{N} + |\dot{\phi}|_\infty |r_i^N - \bar{r}_i|$$

and thus the convergence in (5.13) is insured by the convergence $r_i^N \to \bar{r}_i$, with the order given by $\frac{1}{N}$ if $r_i^N \to \bar{r}_i$ is of this order.

Finally, in considering (5.14) we note that

$$\sum_{j=1}^{N} K_j^N(\alpha^N)\phi_j^N = \sum_{j=1}^{N}\int_{t_j^N}^{t_{j-1}^N} K(\alpha^N,\theta)\phi_j^N d\theta$$

$$= \int_{-r_\nu^N}^{0} K(\alpha^N,\theta)\sum_{j=1}^{N}\phi_j^N\chi_j^N(\theta)\,d\theta$$

and hence

$$\Delta^N \equiv \sum_{j=1}^{N} K_j^N(\alpha^N)\phi_j^N - \int_{-\bar{r}_\nu}^{0} K(\bar{\alpha},\sigma)\phi(\sigma)\,d\sigma$$

$$= \int_{-r_\nu^N}^{0} K(\alpha^N,\sigma)[\Sigma\ \phi_j^N\chi_j^N(\sigma) - \phi(\sigma)]d\sigma + \int_{-r_\nu^N}^{0} K(\alpha^N,\sigma)\phi(\sigma)d\sigma - \int_{-\bar{r}_\nu}^{0} K(\bar{\alpha},\sigma)\phi(\sigma)d\sigma.$$

We thus find

$$|\Delta^N| \le \left(\int_{-r_\nu^N}^{0} |K(\alpha^N,\sigma)|^2 d\sigma\right)^{1/2}\left(\int_{-r_\nu^N}^{0} |\Sigma\ \phi_j^N\chi_j^N - \phi|^2\right)^{1/2}$$

$$+ \left|\int_{-r_\nu^N}^{0} K(\alpha^N,\sigma)\phi(\sigma)d\sigma - \int_{-\bar{r}_\nu}^{0} K(\bar{\alpha},\sigma)\phi(\sigma)d\sigma\right|.$$

The first term is $O(\frac{1}{N})$ by (5.16) while standard estimates on the second term yield that it $\to 0$ since $K(\alpha^N,\cdot) \to K(\bar{\alpha},\cdot)$ in L_2 and $r_\nu^N \to \bar{r}_\nu$. The order of convergence of the second term depends on that of these latter two. If $|r_\nu^N - \bar{r}_\nu| = O(\frac{1}{N})$ and if the convergence $K(\alpha^N,\cdot) \to K(\bar{\alpha},\cdot)$ in L_2 is $O(\frac{1}{N})$, then Δ^N is

$O(\frac{1}{N})$ also.

In summary, we have established (4.7), (4.8) and (4.9) for the averaging approximation I.D. scheme. In doing so we have also shown that under certain circumstances, the convergence in (4.9)-(ii) is $O(\frac{1}{N})$. In particular, if in $q^N \to \bar{q}$ we have

(5.17) (a) The convergence $A_i(\alpha^N) \to A_i(\bar{\alpha})$ is $O(\frac{1}{N})$,

(b) The convergence $K(\alpha^N,\cdot) \to K(\bar{\alpha},\cdot)$ in $L_2(-r_\nu^N,0)$ is $O(\frac{1}{N})$,

(c) $r_i^N \to \bar{r}_i$ is $O(\frac{1}{N})$, $i = 1,2,\ldots,\nu$,

(d) $z = (\phi(0),\phi)$, $\phi \in C^1$ on $[-r,0]$, $\dot{\phi}(0) = 0$,

then $|\mathscr{A}_N(q^N)P_N(\phi(0),\phi) - P_N\mathscr{A}(\bar{q})(\phi(0),\phi)|_N$ is $O(\frac{1}{N})$.

Remark 5.1. We remark that the conditions (5.17a-c) clearly are not conditions that can be verified a priori when using the averaging scheme in practice. These error estimates merely provide information as to how well the scheme might perform when applied to specific I.D. problems. Note that the particular method (maximum likelihood estimator, least squares, etc.) chosen for determining q^N will obviously affect the rates of convergence in (a)-(c) above. Finally, if one drops the condition $\dot{\phi}(0) = 0$ form (5.17d) but retains all other conditions in (5.17), one finds the order in (4.9)-(ii) is only $1/\sqrt{N}$.

<u>Remark 5.2</u>. Recalling the order estimates on (3.2) given in Theorem 3.2, we observe that one can easily find sets $\mathscr{B}$ and $\mathscr{B}_1$ to satisfy the hypothesis of that theorem in case of the averaging based scheme. For example, to insure convergence of order $1/\sqrt{N}$, one can choose the sets $\mathscr{B} = \mathscr{D}(\mathscr{A}^2(\bar{q}))$ and $\mathscr{B}_1 = \mathscr{D}(\mathscr{A}^3(\bar{q}))$ $= \{(\phi(0),\phi) | \ \phi \in W_2^{(3)}(-r,0),\ \dot{\phi}(0) = L(\phi),\ \ddot{\phi}(0) = L(\dot{\phi})\}$. Then one can, under the assumptions (5.17a-c), without difficulty argue the claimed order results.

6. Spline based approximation schemes

We discuss in this section an identification scheme based on spline approximations. While we shall present the details for a scheme based on first order splines, arbitrary order spline approximations may be utilized in a similar manner with only slight modifications in the arguments indicated below (see the theory developed in [10], on which all of our discussions here are based).

Given $q^N = (\alpha^N, r_1^N, \ldots, r_\nu^N) \to \overline{q} = (\overline{\alpha}, \overline{r}_1, \ldots, \overline{r}_\nu)$ as we have hypothesized previously, we partition each of the subintervals $[-r_k^N, -r_{k-1}^N]$, $k = 1,2,\ldots,\nu$, into N equal subintervals to define the partition $\{t_j^N\}_{j=1}^{\nu N}$ of $[-r_\nu^N, 0]$ with

$$(6.1) \qquad t_j^N \equiv -(j-(k-1)N)(r_k^N - r_{k-1}^N)/N + r_{k-1}^N,$$

$j = (k-1)N,\ldots,kN$, $k = 1,2,\ldots,\nu$. We then define the finite dimensional subspace $X_N \subset Z_N$ by

$$X_N = \{(\phi(0),\phi) \mid \phi \text{ is a first order spline with knots at } \{t_j^N\}\}.$$

We define the weighting function g^N by

$$g^N(\theta) = \begin{cases} 1 & -r_\nu^N \le \theta < -r_{\nu-1}^N, \\ 2 & -r_{\nu-1}^N \le \theta < r_{\nu-2}^N, \\ \vdots & \\ \nu-1 & -r_2^N \le \theta < -r_1^N \\ \nu & -r_1^N \le \theta < 0, \end{cases}$$

and, as in §5, denote by $Z_N(g^N)$ and $X_N(g^N)$ the spaces Z_N and X_N endowed with the equivalent topology generated by the weighted inner product (5.7). We then define $\pi_N: Z_N \to X_N$ (equivalently $\pi_N: Z_N(g^N) \to X_N(g^N)$) as the orthogonal projection of $Z_N(g^N)$ onto $X_N(g^N)$. Then (see [10 , p. 509]) for ψ in Z_N, we have $\pi_N\psi = \hat{\psi}^N$ where $\hat{\psi}^N$ is the solution of the problem of minimizing $|\hat{\psi}-\psi|_{Z_N(g^N)}$ over $\hat{\psi} \in X_N$. The operator $P_N: Z \to X_N$ is defined as before by $P_N = \pi_N \mathscr{I}_N$.

We adopt the following notation. For any function ϕ that is defined pointwise on $[-r_\nu^N,0]$, we write $\hat{\phi} = (\phi(0),\phi)$ and $\hat{\phi}_I^N = (\phi_I^N(0),\phi_I^N)$ where ϕ_I^N is the interpolating spline (with knots at $\{t_j^N\}_{j=1}^{\nu N}$) for ϕ on $[-r_\nu^N,0]$. For the projections π_N defined above we shall write $\pi_N\hat{\phi} = \pi_N(\phi(0),\phi) = \hat{\phi}^N = (\phi^N(0),\phi^N)$.

For any $q^N \in Q$, we define the operator $\mathscr{A}(q^N)$: $\mathscr{D}(\mathscr{A}(q^N)) \subset Z_N \to Z_N$, where $\mathscr{D}(\mathscr{A}(q^N))$ $= \{(\phi(0),\phi) \in Z_N | \phi \in W_2^{(1)}(-r_\nu^N,0)\}$, by

$$\mathscr{A}(q^N)(\phi(0),\phi) = (L(q^N)\phi, D\phi).$$

More generally, for any $q \in Q$, we can define $\mathscr{A}(q)$: $\mathscr{D} \subset Z \to Z$ by

$$\mathscr{A}(q)(\phi(0),\phi) = (L(q)\phi, D\phi).$$

Note that in this latter case $D\phi$ is defined on $[-r,0]$, while in the former $D\phi$ is defined on $[-r_\nu^N,0]$. However, in both cases the operators are essentially the same and in the discussions below we shall abuse notation and speak of $\mathscr{A}$ as an operator defined either in Z_N or Z, depending on the context. We note that $X_N \subset \mathscr{D}(\mathscr{A}(q^N))$ so that $\mathscr{A}(q^N)$ is defined on all of X_N.

With the definitions above, we have immediately from Lemma 2.3 of [10] that $\mathscr{A}(q^N)$ satisfies

(6.2) $$\langle \mathscr{A}(q^N)z,z\rangle_{g^N} \leq \omega(q^N)|z|^2_{g^N}, \quad z \in \mathscr{D}(\mathscr{A}(q^N)),$$

where

$$\omega(q^N) \equiv \frac{\nu+1}{2} + |A_0(\alpha^N)| + \frac{1}{2}\sum_{i=1}^{\nu}|A_i(\alpha^N)|^2 + \frac{1}{2}\int_{-r_\nu^N}^{0}|K(\alpha^N,\theta)|^2 d\theta .$$

We next define $\mathscr{A}_N(q^N)\colon X_N \to X_N$ by

(6.3) $$\mathscr{A}_N(q^N) = \pi_N \mathscr{A}(q^N)\pi_N .$$

In view of (6.2) and the fact that $\pi_N x = x$ for any $x \in X_N$, we find for every $x \in X_N$

$$\begin{aligned}\langle \mathscr{A}_N(q^N)x,x\rangle_{g^N} &= \langle \pi_N \mathscr{A}(q^N)x,x\rangle \\ &= \langle \mathscr{A}(q^N)x,x\rangle \leq \omega(q^N)|x|^2_{g^N}.\end{aligned}$$

As noted in §5, there exists β such that $\omega(q^N) \leq \beta$ for all N and thus $\mathscr{A}_N(q^N) \in G(\tilde{M},\beta)$ on $X_N(g^N)$ and hence $\mathscr{A}_N(q^N) \in G(M,\beta)$ on X_N for some M independent of N. Similar arguments establish that $\mathscr{A}(\bar{q}) \in G(M,\beta)$ on Z for M,β appropriately chosen and it thus follows that condition (4.8) is satisfied by the approximations (6.3).

We turn next to the consistency condition (4.9) and define $\mathscr{D}_1 \equiv \mathscr{D}(\mathscr{A}^3(\bar{q}))$. This set is dense in Z and it follows at once that $R_\lambda(\mathscr{A}(\bar{q}))\mathscr{D}_1 \subset \mathscr{D}_1$ so that (4.9)-(i) is satisfied. For $z = (\phi(0),\phi) \in \mathscr{D}_1$ we have

$$\mathscr{A}_N(q^N)P_N z = \pi_N(L(q^N)\hat{\phi}^N, D\hat{\phi}^N)$$

where $\hat{\phi}^N = P_N(\phi(0),\phi) = \pi_N \mathscr{I}_N(\phi(0),\phi)$, while

$$P_N\mathscr{A}(\bar{q})z = \pi_N \mathscr{I}_N(L(\bar{q})\phi, D\phi).$$

It thus follows that for $z \in \mathscr{D}_1$

$$(6.4) \quad |\mathscr{A}_N P_N z - P_N \mathscr{A} z|_N = |\pi_N(L(q^N)\hat{\phi}^N - L(\bar{q})\phi,\ D(\hat{\phi}^N - \phi))|_N$$

where we now understand that $D(\hat{\phi}^N-\phi)$ is to be taken as a function on $[-r_\nu^N, 0]$. Recalling that π_N is the orthogonal projection of $Z_N(g^N)$ onto $X_N(g^N)$ and that the norms of $Z_N(g^N)$ and Z_N are equivalent (with constants independent of N), i.e. $|\pi_N z|_{g^N} \le |z|_{g^N} \le m|z|_N$, we see that (6.4) allows us to establish condition (4.9)-(ii) by verifying

$$(6.5) \qquad |D(\hat{\phi}^N-\phi)| \to 0 \quad \text{as} \quad N \to \infty$$

and

$$(6.6) \qquad |L(q^N)\phi^N - L(\bar{q})\phi|_{R^n} \to 0 \quad \text{as} \quad N \to \infty,$$

where in (6.5) the norm can be that of $L_2(-r_\nu^N,0)$ or L_2 with the weighting function (see (5.7)) g^N.

To show (6.5) and (6.6) we shall make use of standard estimates from the theory of spline approximations. Specifically from Theorem 2.5 of [24] we find, upon considering the interval $[t_{kN}^N, t_{(k-1)N}^N]$, $k = 1,2,\ldots,\nu$, which has mesh size $h = (r_k^M - r_{k-1}^N)/N$,

$$(6.7) \qquad \int_{t_{kN}^N}^{t_{(k-1)N}^N} |D(\phi-\phi_I^N)|^2 \le \frac{1}{\pi^2}\left\{\frac{r_k^N - r_{k-1}^N}{N}\right\}^2 \int_{t_{kN}^N}^{t_{(k-1)N}^N} |D^2\phi|^2$$

and

$$(6.8) \qquad \int_{t_{kN}^N}^{t_{(k-1)N}^N} |\phi-\phi_I^N|^2 \le \frac{1}{\pi^4}\left\{\frac{r_k^N - r_{k-1}^N}{N}\right\}^4 \int_{t_{kN}^N}^{t_{(k-1)N}^N} |D^2\phi|^2.$$

Here ϕ_I^N is the interpolating spline for $\phi \in C^2[-r,0]$ with knots at $\{t_j^N\}$. Denoting by $|\cdot|_{2,N}$ the norm in $L_2(-r_\nu^N,0)$, we deduce from (6.7) and (6.8) the estimates

$$(6.9) \qquad |D(\phi-\phi_I^N)|_{2,N} \le \frac{1}{\pi}\{\max_j |r_j^N - r_{j-1}^N|\}\frac{1}{N}\,|D^2\phi|_{2,N}$$

and

$$(6.10) \qquad |\phi-\phi_I^N|_{2,N} \leq \frac{1}{\pi^2}\{\max_j |r_j^N - r_{j-1}^N|\}^2 \frac{1}{N^2} |D^2\phi|_{2,N}.$$

Denoting by $|\cdot|_{2,n,g^N}$ the weighted norm in $L_2(-r_\nu^N,0)$, we easily argue for $\hat{\phi} = (\phi(0),\phi) \in \mathscr{D}_1$ (using the minimality properties associated with π_N)

$$|\phi^N-\phi|_{2,N} \leq |\pi_N\hat{\phi}-\hat{\phi}|_N = |\hat{\phi}^N-\hat{\phi}|_N \leq |\hat{\phi}^N-\hat{\phi}|_{g^N}$$

$$\leq |\hat{\phi}_I^N-\hat{\phi}|_{g^N} = |\phi_I^N-\phi|_{2,N,g^N} \leq \sqrt{\nu}|\phi_I^N-\phi|_{2,N}.$$

From (6.10) we observe that this last quantity is $O(\frac{1}{N^2})$ and hence so is $|\pi_N\hat{\phi} \;\; \hat{\phi}|_N$. It follows immediately that $|\phi^N(0)-\phi(0)|_{R^n}$ is $O(\frac{1}{N^2})$ also.

We remark that we have shown that $|\pi_N\hat{\phi}-\hat{\phi}|_N \to 0$ whenever $\hat{\phi} \in \mathscr{D}_1$. The density of $\mathscr{D}_1$ in Z and the boundedness of $\{\pi_N\}$ thus imply this convergence $(|\pi_N z-z|_N \to 0)$ for all $z \in Z$ and the condition $\pi^0(P_N z) \to \pi^0 z$ for $z \in Z$ of (4.7) is satisfied.

We next consider the inequality

$$(6.11) \qquad |D(\phi^N-\phi)|_{2,N} \leq |D(\phi^N-\phi_I^N)|_{2,N} + |D(\phi_I^N-\phi)|_{2,N}$$

and observe that the second term is $O(\frac{1}{N})$ by (6.9). We employ the Schmidt inequality [24] to estimate the first term. Since both ϕ^N and ϕ_I^N are linear on each subinterval $[t_j^N,t_{j-1}^N]$ we have

$$(6.12)\qquad \int_{-r_\nu^N}^{0} |D(\phi^N-\phi_I^N)|^2 = \sum_{j=1}^{\nu N} \int_{t_j^N}^{t_{j-1}^N} |D(\phi^N-\phi_I^N)|^2$$

$$\le \sum_{j=1}^{\nu N} \frac{\mathscr{K}}{(t_{j-1}^N-t_j^N)^2} \int_{t_j^N}^{t_{j-1}^N} |\phi^N-\phi_I^N|^2$$

$$= \sum_{k=1}^{\nu} \frac{\mathscr{K}}{[(r_k^N-r_{k-1}^N)/N]^2} \int_{-r_k^N}^{-r_{k-1}^N} |\phi^N-\phi_I^N|^2$$

$$\le \sum_{k=1}^{\nu} \frac{\mathscr{K}}{[(r_k^N-r_{k-1}^N)/N]^2} \left\{ \int_{-r_k^N}^{-r_{k-1}^N} |\phi^N-\phi|^2 + \int_{-r_k^N}^{-r_{k-1}^N} |\phi-\phi_I^N|^2 \right\}$$

$$= T_1^N + T_2^N.$$

Using (6.8) we obtain the estimate for T_2^N

$$T_2^N \le \sum_{k=1}^{\nu} \frac{\mathscr{K}}{\pi^4} \left\{ \frac{r_k^N-r_{k-1}^N}{N} \right\}^2 \int_{-r_k^N}^{-r_{k-1}^N} |D^2\phi|^2 \le \frac{\mathscr{K}}{\pi^4} \left(\frac{r}{N}\right)^2 |D^2\phi|_{2,N}.$$

To obtain the desired estimate on T_1^N we need an additional assumption on $\overline{q} = (\overline{\alpha},\overline{r}_1,\ldots,\overline{r}_\nu)$. Specifically we assume:

(6.13) There exist $\delta > 0$ such that $|\overline{r}_k-\overline{r}_{k-1}| \ge \delta$, $k = 1,2,\ldots,\nu$.

With the assumption we find (for N sufficiently large)

$$T_1^N \leq \frac{\mathscr{K} N^2}{(\delta/2)^2} \int_{-r_\nu^N}^{0} |\phi^N - \phi|^2 = \frac{4\mathscr{K} N^2}{\delta^2} |\phi^N - \phi|^2_{2,N}.$$

But our arguments above revealed that $|\phi^N - \phi|_{2,N}$ is $O(\frac{1}{N^2})$ and thus T_1^N, like T_2^N, is $O(\frac{1}{N^2})$. It follows from (6.12) that the first term in (6.11) is $O(\frac{1}{N})$. We have thus, under the additional assumption (6.13), established (6.5).

Finally, we observe that, for $-r_\nu^N \leq \theta \leq 0$,

$$\phi^N(\theta) = \phi^N(0) + \int_0^\theta D\phi^N$$
$$\phi(\theta) = \varphi(0) + \int_0^\theta D\phi$$

and thus

$$|\phi^N(\theta) - \phi(\theta)| \leq |\phi^N(0) - \varphi(0)| + \int_{-r_\nu^N}^{0} |D\varphi^N - D\phi|$$
$$\leq |\phi^N(0) - \phi(0)| + \sqrt{r}\,|D(\phi^N - \phi)|_{2,N}.$$

But these last two terms are $O(\frac{1}{N})$, uniformly in θ. It follows that $|\phi^N(-r_i^N) - \phi(-r_i^N)|$ is $O(\frac{1}{N})$. Since ϕ is continuous and $q^N \to \bar{q}$ we find that $L(q^N)\phi^N \to L(\bar{q})\phi$ and thus (6.6) obtains.

Summarizing, we have shown that (4.7), (4.8), (4.9) (where π_N is now the projection of $Z_N(g^N)$ onto $X_N(g^N)$) hold for the first order spline based scheme defined by the operators in (6.3) under the assumption (6.13). Furthermore, if one inspects carefully the estimates given above, one finds that under the hypothesis

(5.17a-c), the convergence in (4.9)(ii) is $O(\frac{1}{N})$.

Remark 6.1. In the above estimates we had chosen $\mathscr{D}_1 = \mathscr{D}(\mathscr{A}^3(\bar{q}))$ so that ϕ in (6.4) and the subsequent arguments was in $W_2^{(3)}(-r,0)$. To apply the needed estimates (e.g., Theorem 2.5 of [24]) and make the arguments above, it is actually sufficient to have ϕ in $W_2^{(2)}(-r,0)$ (see Theorem 21 of [25]). We thus could have just as easily chosen $\mathscr{D}_1 = \mathscr{D}(\mathscr{A}^2(\bar{q}))$ and arrived at the conclusions above, including the convergence rates obtained.

Remark 6.2. In light of the above remark, we may, in order to obtain that the approximating semigroups converge like $O(\frac{1}{N})$ for the scheme developed here, choose $\mathscr{B} = \mathscr{D}(\mathscr{A}^2(\bar{q}))$ and $\mathscr{B}_1 = \mathscr{D}(\mathscr{A}^3(\bar{q}))$ in Theorem 3.2. Under assumptions (6.13) and (5.17a-c), one then can readily verify that the hypotheses of Theorem 3.2 are satisfied by the spline-based approximations.

7. Numerical results

In this section we present a brief summary of some numerical results for the identification problem obtained using the approximation schemes (AVE and SPLINE) outlined in the two previous sections. For a more detailed discussion of the numerical performance of the AVE and SPLINE schemes in identification and control problems, the reader can consult [8] where numerous examples, error analyses, etc. are presented. The summary given here, taken with the extensive numerical tests reported in [8], support our claims of efficacy and practical usefulness for these methods.

In order to generate the data for testing the algorithms we select a "true" set of parameters $\gamma^* = (\eta^*,\phi^*,\alpha^*,r^*)$ (we take $\nu = 1$ and $r_1 = r$) and a control u and use the method of steps [14] to solve for x on the interval $[0,T]$. In all of the examples presented below "data" was generated using $r^* = 1$ and $u = u_\ell$, where u_ℓ is the unit step at $t = \ell$ defined by

$$u_\ell(t) = \begin{cases} 0 & t < \ell, \\ 1 & \ell \leq t, \end{cases}$$

and $0 < \ell < 1$. The final time of $T = 2$ was used. The observations $\hat{y}_i = y(t_i)$ were generated at 101 equally spaced time steps on $[0,T]$. It is possible to add noise to the "data" to produce "noisy observations" $\hat{y}(t) = y(t) + \nu(t)$, where, for example, $\nu(t) = \mathrm{col}(\nu_1(t), \ldots, \nu_k(t))$ is a computer-simulated vector of normal random variables $\nu_i(t)$, each with zero mean and preset standard variation. This was done for some of the examples in [8] (we do not do it here) where one again finds that the algorithms perform quite well.

For each fixed N, the approximation problem ($\mathcal{P}_N$) was solved using a maximum likelihood estimator (MLE). The resulting solutions are denoted $\bar{\gamma}_A^N$ and $\bar{\gamma}_S^N$ for the AVE and SPLINE schemes respectively. Since the MLE is an iterative procedure it is necessary to supply a startup value (i.e. an initial guess) for the parameters γ_A^N or γ_S^N. If β denotes an unknown parameter to be estimated (e.g. $\beta = \alpha$ or $\beta = r$), then $\beta^{N,I}$ will denote the estimate for $\bar{\beta}^N$ obtained after I iterations of the MLE applied to problem ($\mathcal{P}_N$). The startup value is denoted by $\beta^{N,0}$.

<u>Example 7.1</u>

In this example we seek to estimate the initial data $(\eta,\phi) \in R \times L_2(-1,0)$ and the coefficient of the delayed term in a simple scalar equation. The system is described by the equation

$$\dot{x}(t) = .05\, x(t) + a_1 x(t-1) + u_{.1}(t),$$

with (unknown) initial data

$$x(0) = \eta, \qquad x_0(s) = \phi(s), \qquad -1 \leq s < 0,$$

and output

$$y(t) = x(t).$$

Data was generated as described previously using the true values $\eta^* = 1$, $\phi^* \equiv 1$, $a_1^* = -4.0$. For each N = 2, 4, 8, 16 and 32, the approximating problem $(\mathcal{P}_N)$ was formulated as discussed in section 4. Thus, for AVE we seek the "parameter"

$$\bar{\gamma}_A^N = (\eta, \phi_1^N, \phi_2^N, \ldots, \phi_N^N, a_1),$$

where $(\eta, \phi_1^N, \phi_2^N, \ldots, \phi_N^N)$ are coordinates of the AVE projection of the initial data. Similarly, for SPLINE we seek the "parameter"

$$\bar{\gamma}_S^N = (\xi_0^N, \xi_1^N, \ldots, \xi_N^N, a_1),$$

where $(\xi_0^N, \xi_1^N, \ldots, \xi_N^N)$ are coordinates for the SPLINE projection of the initial data. The "start-up" for $(\eta,\phi) \in R \times L_2(-1,0)$ was taken as the zero initial data (0,0), whereas the "start-up" for a_1 was chosen as $a_1^{N,0} = -3.0$. Table 7.1.1 provides an overview of the numerical findings. Because the initial data is in $R \times L_2(-1,0)$ we have only displayed the Z-norm of the error and the estimated value for a_1^N. The comparison of the two schemes is quite striking; in particular, note the relative accuracies in estimating the initial data. Compared in Figure 7.1.1 are graphs of the true initial data and the corresponding estimates produced by AVE and SPLINE for N = 4. It is apparent that (at least for the chosen "start-up" values) the SPLINE procedure readily finds good estimates for the parameters, while the AVE scheme has considerable difficulty.

It is of some interest to compare the sequence of data fits generated as the MLE iteration procedure evolves. Figures 7.1.2, 7.1.3 and 7.1.4 show the data matches from the AVE algorithm (with N = 8) for MLE iterations 0, 4 and 9, respectively. From the match at iteration 4 (Figure 7.1.3) it might be deduced that the AVE scheme is in trouble. However, at iteration 9 the fit is quite good and Figure 7.1.4 does not give any hint of the poor values of the parameters indicated in Table 7.1.1.

Figures 7.1.5, 7.1.6 and 7.1.7 illustrate the SPLINE matches at iterations 0, 4 and 9, respectively. Again the iteration 4 matches indicate some difficulty while by iteration 9 the match is quite good. It happens that the SPLINE estimates of the parameters are excellent.

Although one cannot be certain, for the AVE scheme it does appear that the MLE procedure is converging to a local minimum of J^N. We suspect, however, that the problem $(\mathscr{P}_N)$ suffers a lack of identifiability. (See [8] for a further discussion of this matter.) The problem $(\mathscr{P}_N)$ for SPLINE seems to be much better behaved.

In order to investigate further identifiability for problems with unknown initial data, we made further computations for this example using the same dynamics, changing only the initial data to

$$\eta = 1, \quad \phi(s) = 1 + s, \quad -1 \leq s < 0.$$

Using the same start-ups as above, we found that SPLINE converged for all N values, whereas AVE never did. Results are summarized in Table 7.1.2.

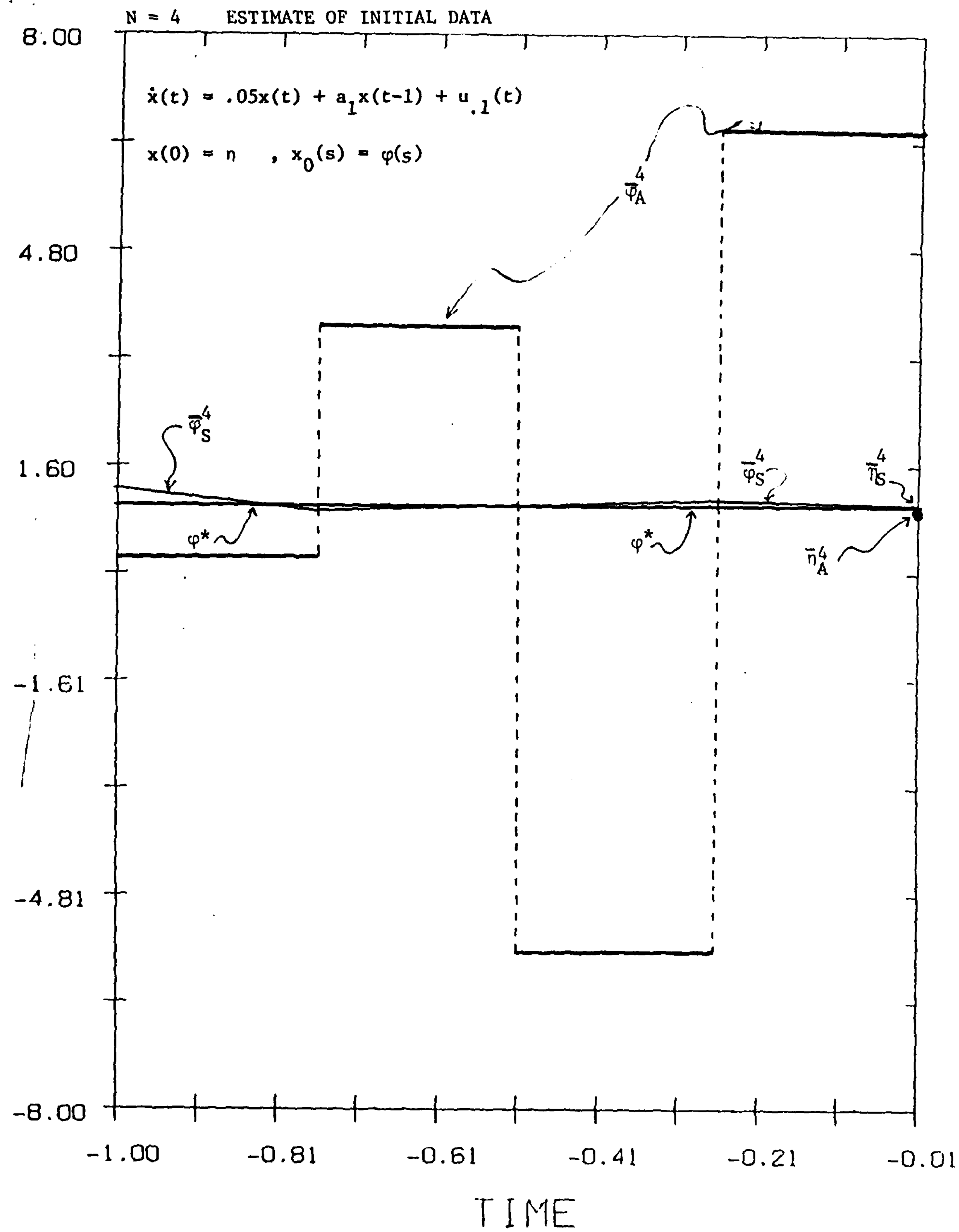
AVE			SPLINE		
N	$\bar{a}_1^N$	$\lvert z^*(0)-\bar{z}^N(0)\rvert$	N	$\bar{a}_1^N$	$\lvert z^*(0)-\bar{z}^N(0)\rvert$
2	-4.4103	2.08	2	-4.4382	.1595
4	-4.9924	4.53	4	-3.9381	.0867
8	-4.2651	41.76	8	-4.0031	.0287
16	did not converge		16	-4.0031	.0201
32	did not converge		32	-4.0001	.0386

TABLE 7.1.1

AVE			SPLINE		
N	$\bar{a}_1^N$	$\lvert z^*(0)-\bar{z}^N(0)\rvert$	N	$\bar{a}_1^N$	$\lvert z^*(0)-\bar{z}^N(0)\rvert$
2	did not converge		2	-4.5201	.0563
4			4	-4.0975	.0318
8			8	-4.0282	.0123
16			16	-4.0123	.0193
32			32	-4.0122	.0936

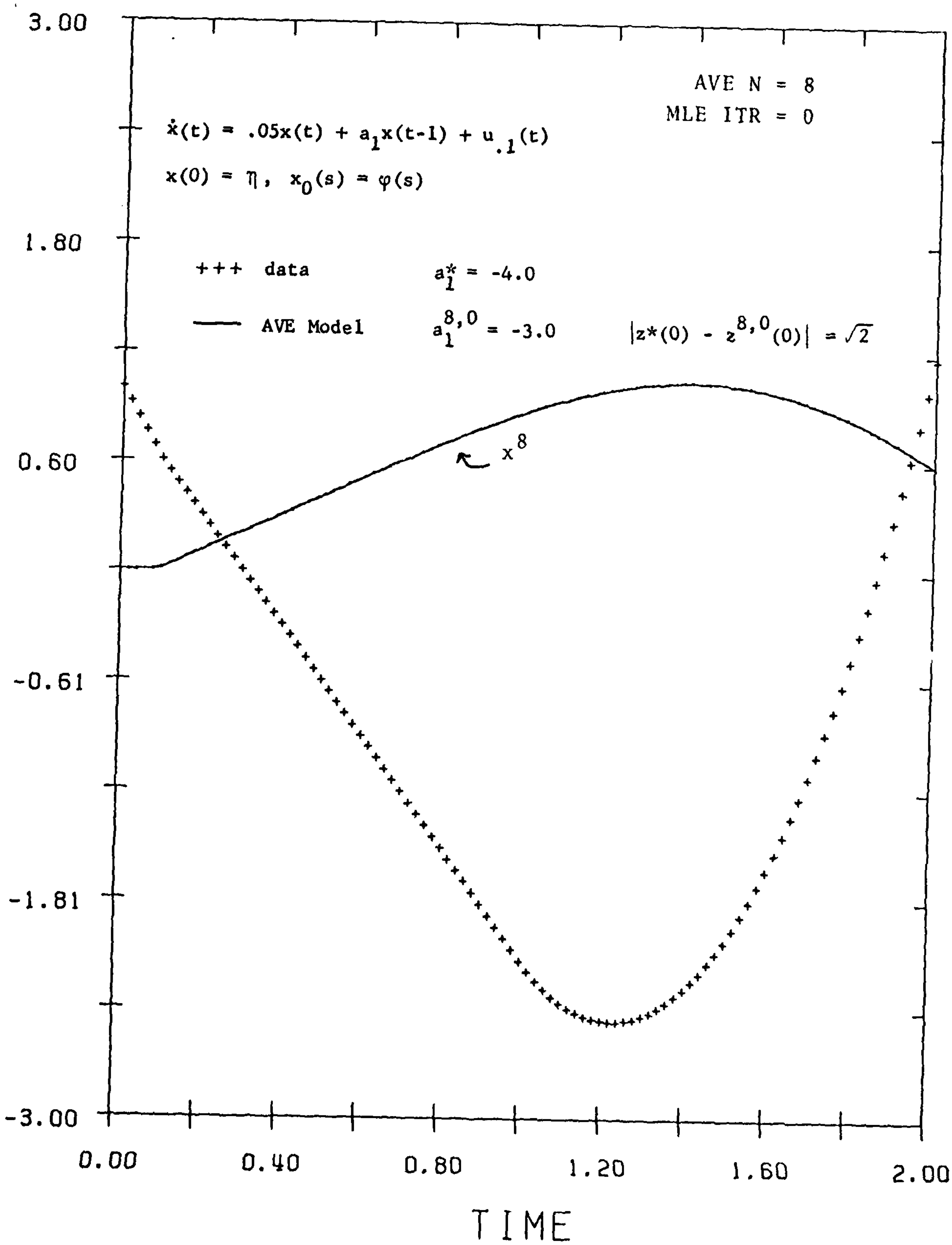
TABLE 7.1.2

(linear initial data)



N = 4 ESTIMATE OF INITIAL DATA
$\dot{x}(t) = .05x(t) + a_1x(t-1) + u_{.1}(t)$
$x(0) = \eta \quad , \; x_0(s) = \varphi(s)$
$\bar{\varphi}_A^4$
$\bar{\varphi}_S^4$
φ^*
$\bar{\eta}_S^4$
$\bar{\eta}_A^4$
8.00
4.80
1.60
-1.61
-4.81
-8.00
-1.00
-0.81
-0.61
-0.41
-0.21
-0.01
TIME

FIGURE 7.1.1



3.00
1.80
0.60
-0.61
-1.81
-3.00
0.00
0.40
0.80
1.20
1.60
2.00
TIME
AVE N = 8
MLE ITR = 0
$\dot{x}(t) = .05x(t) + a_1x(t-1) + u_{.1}(t)$
$x(0) = \eta, \; x_0(s) = \varphi(s)$
+++ data
$a_1^* = -4.0$
AVE Model
$a_1^{8,0} = -3.0$
$|z^*(0) - z^{8,0}(0)| = \sqrt{2}$
x^8

FIGURE 7.1.2

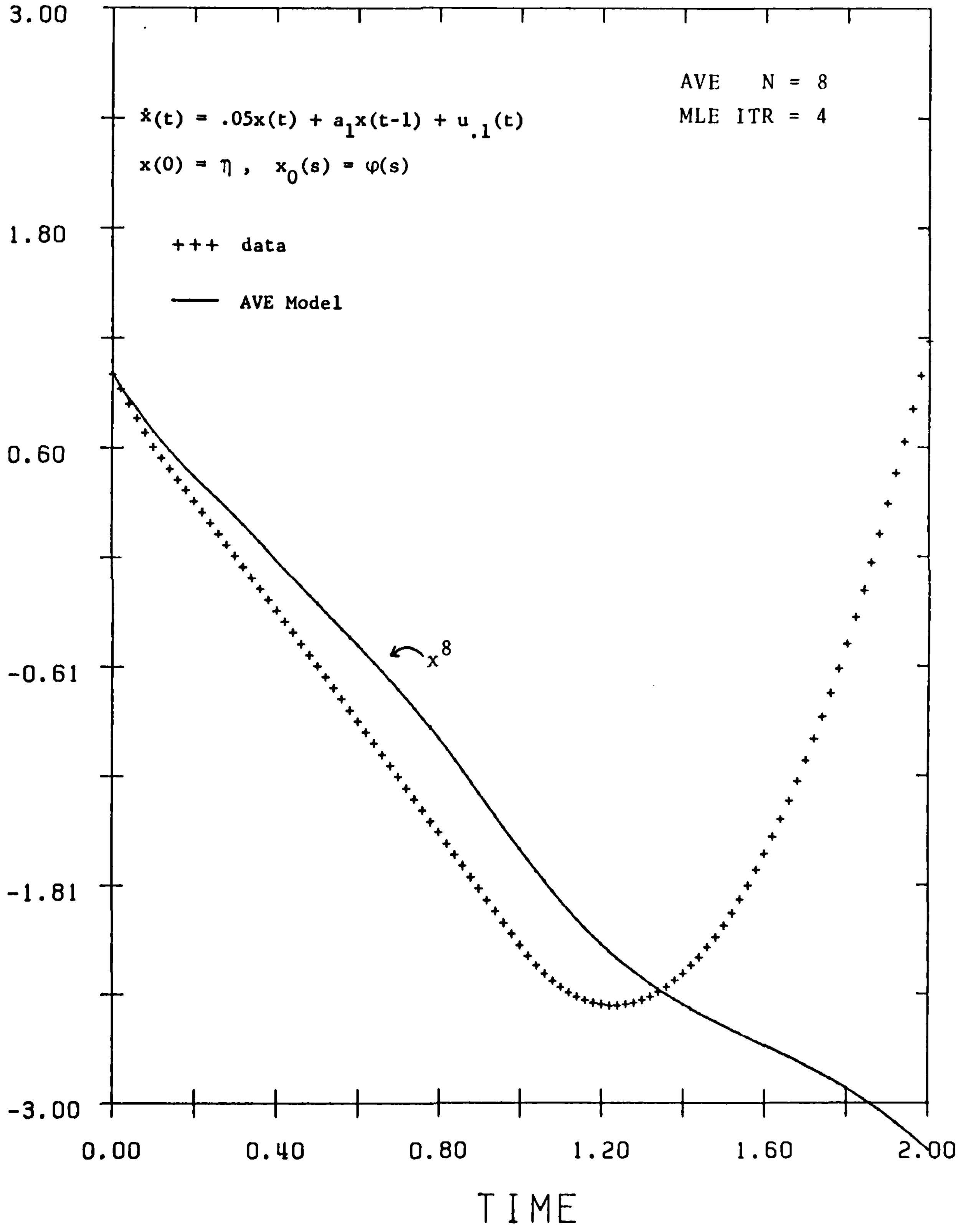

3.00
1.80
0.60
-0.61
-1.81
-3.00
0.00
0.40
0.80
1.20
1.60
2.00
TIME
AVE N = 8
MLE ITR = 4
$\dot{x}(t) = .05x(t) + a_1x(t-1) + u_{.1}(t)$
$x(0) = \eta$, $x_0(s) = \varphi(s)$
+++ data
AVE Model
x^8

FIGURE 7.1.3

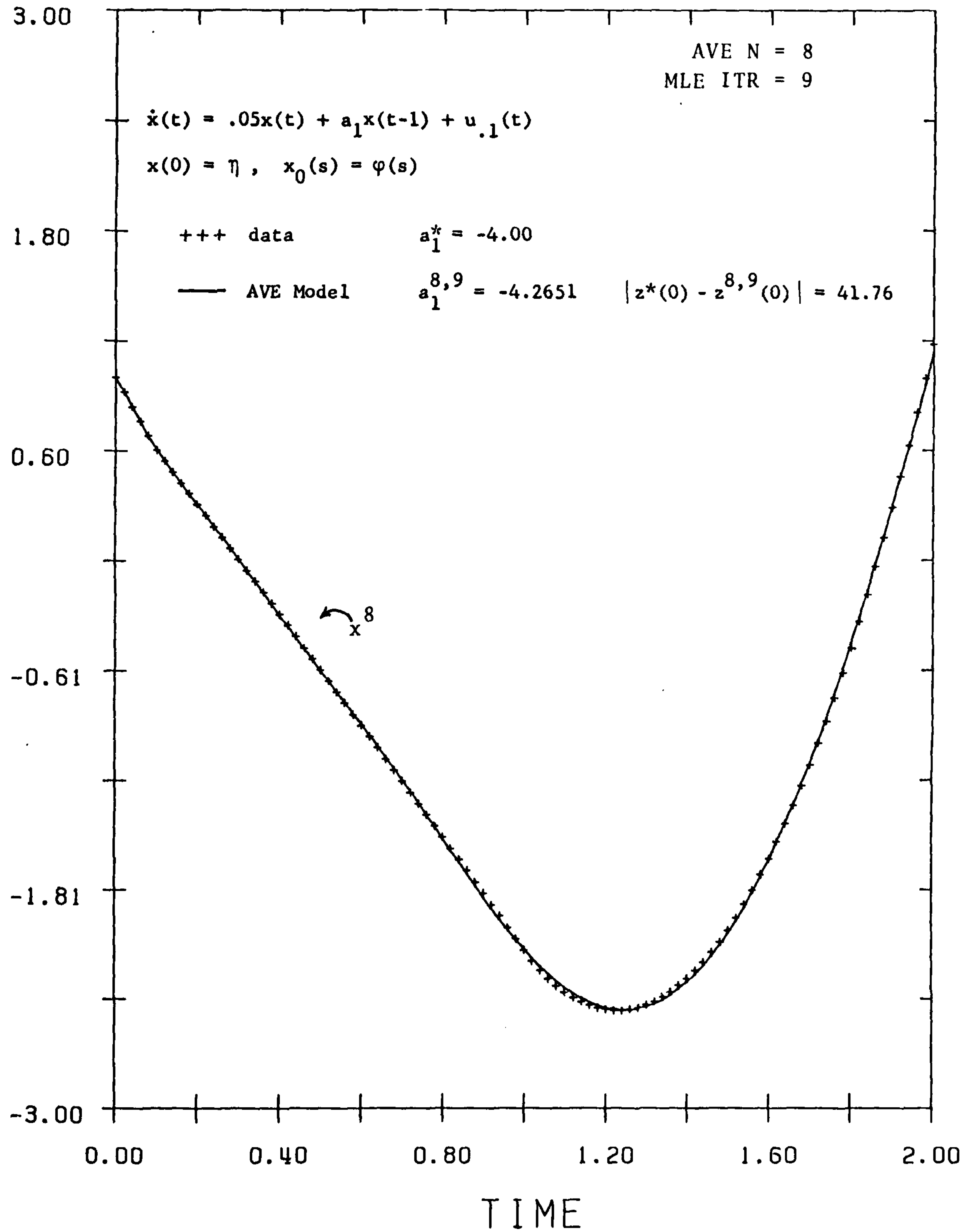
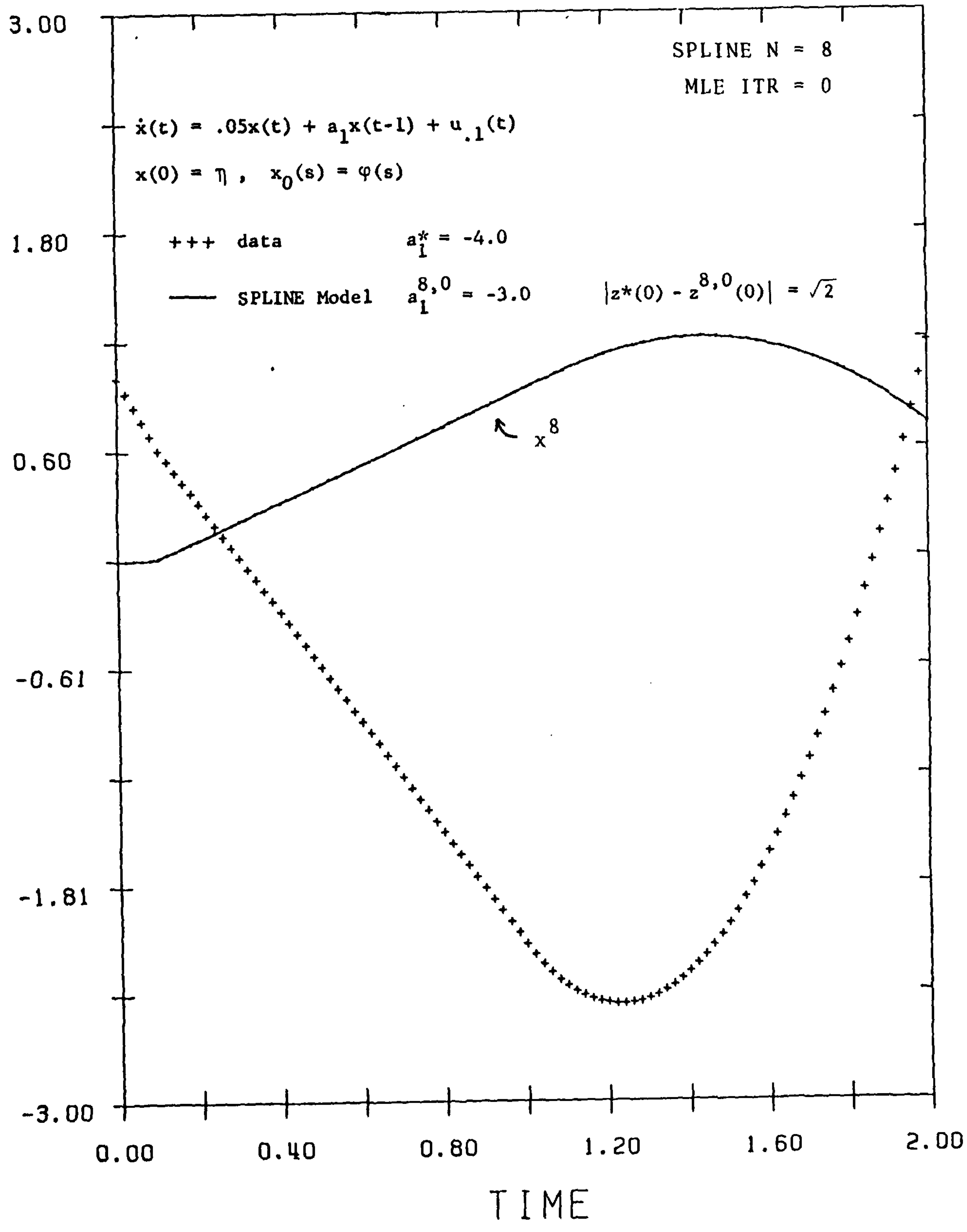

3.00
1.80
0.60
-0.61
-1.81
-3.00
0.00
0.40
0.80
1.20
1.60
2.00
TIME
AVE N = 8
MLE ITR = 9
$\dot{x}(t) = .05x(t) + a_1x(t-1) + u_{.1}(t)$
$x(0) = \eta$, $x_0(s) = \varphi(s)$
+++ data
$a_1^* = -4.00$
AVE Model
$a_1^{8,9} = -4.2651$
$|z^*(0) - z^{8,9}(0)| = 41.76$
x^8

FIGURE 7.1.4



3.00
1.80
0.60
-0.61
-1.81
-3.00
0.00
0.40
0.80
1.20
1.60
2.00
TIME
SPLINE N = 8
MLE ITR = 0
$\dot{x}(t) = .05x(t) + a_1x(t-1) + u_{.1}(t)$
$x(0) = \eta, \quad x_0(s) = \varphi(s)$
+++ data
$a_1^* = -4.0$
SPLINE Model
$a_1^{8,0} = -3.0$
$|z^*(0) - z^{8,0}(0)| = \sqrt{2}$
x^8

FIGURE 7.1.5

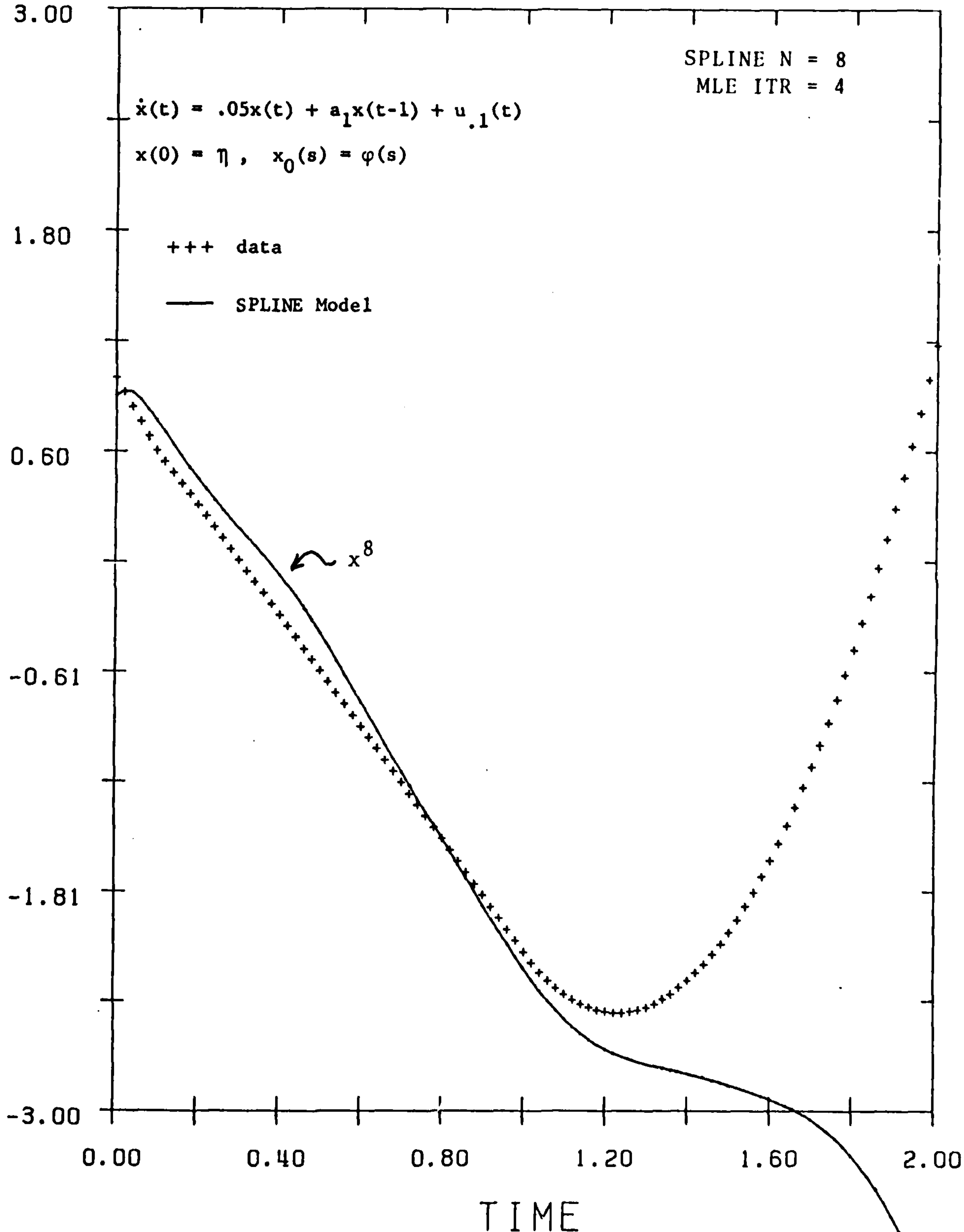
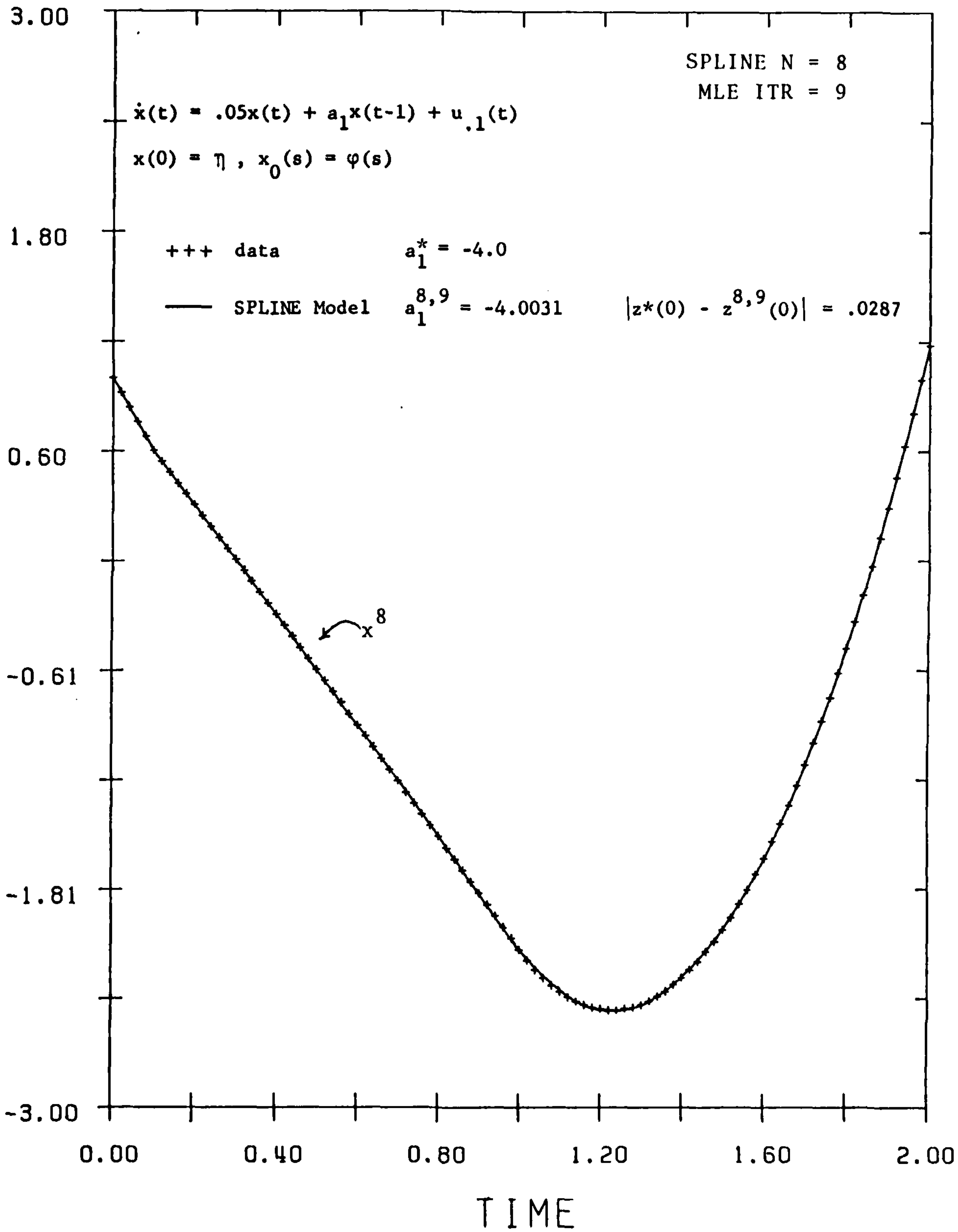

3.00
1.80
0.60
-0.61
-1.81
-3.00
SPLINE N = 8
MLE ITR = 4
$\dot{x}(t) = .05x(t) + a_1x(t-1) + u_{.1}(t)$
$x(0) = \eta, \quad x_0(s) = \varphi(s)$
+++ data
SPLINE Model
x^8
0.00
0.40
0.80
1.20
1.60
2.00
TIME

FIGURE 7.1.6



3.00
1.80
0.60
-0.61
-1.81
-3.00
SPLINE N = 8
MLE ITR = 9
$\dot{x}(t) = .05x(t) + a_1x(t-1) + u_{.1}(t)$
$x(0) = \eta$, $x_0(s) = \varphi(s)$
+++ data
$a_1^* = -4.0$
SPLINE Model
$a_1^{8,9} = -4.0031$
$|z^*(0) - z^{8,9}(0)| = .0287$
x^8
0.00
0.40
0.80
1.20
1.60
2.00
TIME

FIGURE 7.1.7

Example 7.2

We consider an equation with a continuous (a constant function) kernel in which we wish to estimate the kernel, a system coefficient, and the time delay. The model is assumed to be of the form

$$\dot{x}(t) = a_1 x(t-r) + k \int_{-r}^{0} x(t+s)ds + u_{.1}(t),$$

with initial data

$$x_0(s) \equiv 1, \qquad -r \le s \le 0,$$

and output

$$y(t) = x(t).$$

The true parameters $a_1^* = -3.0$, $k^* = -1.0$ and $r^* = 1.0$ were estimated using start-ups of

$$a_1^{N,0} = -3.5, \quad k^{N,0} = -1.5, \qquad r^{N,0} = 1.5.$$

Runs were made for $N = 2, 4, 8$ and 16. The MLE algorithm for the AVE scheme did not converge for $N = 2$ and 4. However, for $N = 8$ and 16 the AVE scheme converged but produced rather poor parameter estimates. The SPLINE scheme converged for each $N = 2, 4, 8, 16$ and for $N \ge 4$ produced good parameter estimates. The numerical results for this problem are summarized in Tables 7.2.1 and 7.2.2, where $e_N \equiv \bar{\gamma}^N - \gamma^*$ is the error.

Figures 7.2.1 through 7.2.4 compare the N = 8 AVE and SPLINE data fits. In particular, Figures 7.2.1 and 7.2.2 show the N = 8 AVE start-up and converged data fits, respectively. Figures 7.2.3 and 7.2.4 show similar results for the SPLINE procedure.

AVE				
N	$\bar{r}^N$	$\bar{k}^N$	$\bar{a}_1^N$	$\lvert e_N \rvert$
2		- did not converge -		
4		- did not converge -		
8	.8802	.2182	-4.1641	2.0657
16	.9383	-.3806	-3.5535	1.2346
γ^* =	1.0000	-1.0000	-3.0000	

Table 7.2.1

SPLINE				
N	$\bar{r}^N$	$\bar{k}^N$	$\bar{a}_1^N$	$\lvert e_N \rvert$
2	.9100	-.4376	-3.4478	1.1002
4	.9896	-1.0087	-3.0580	.0071
8	1.0018	-1.0390	-2.9953	.0455
16	1.0042	-1.0410	-2.9841	.0611
γ^* =	1.0000	-1.0000	-3.0000	

Table 7.2.2

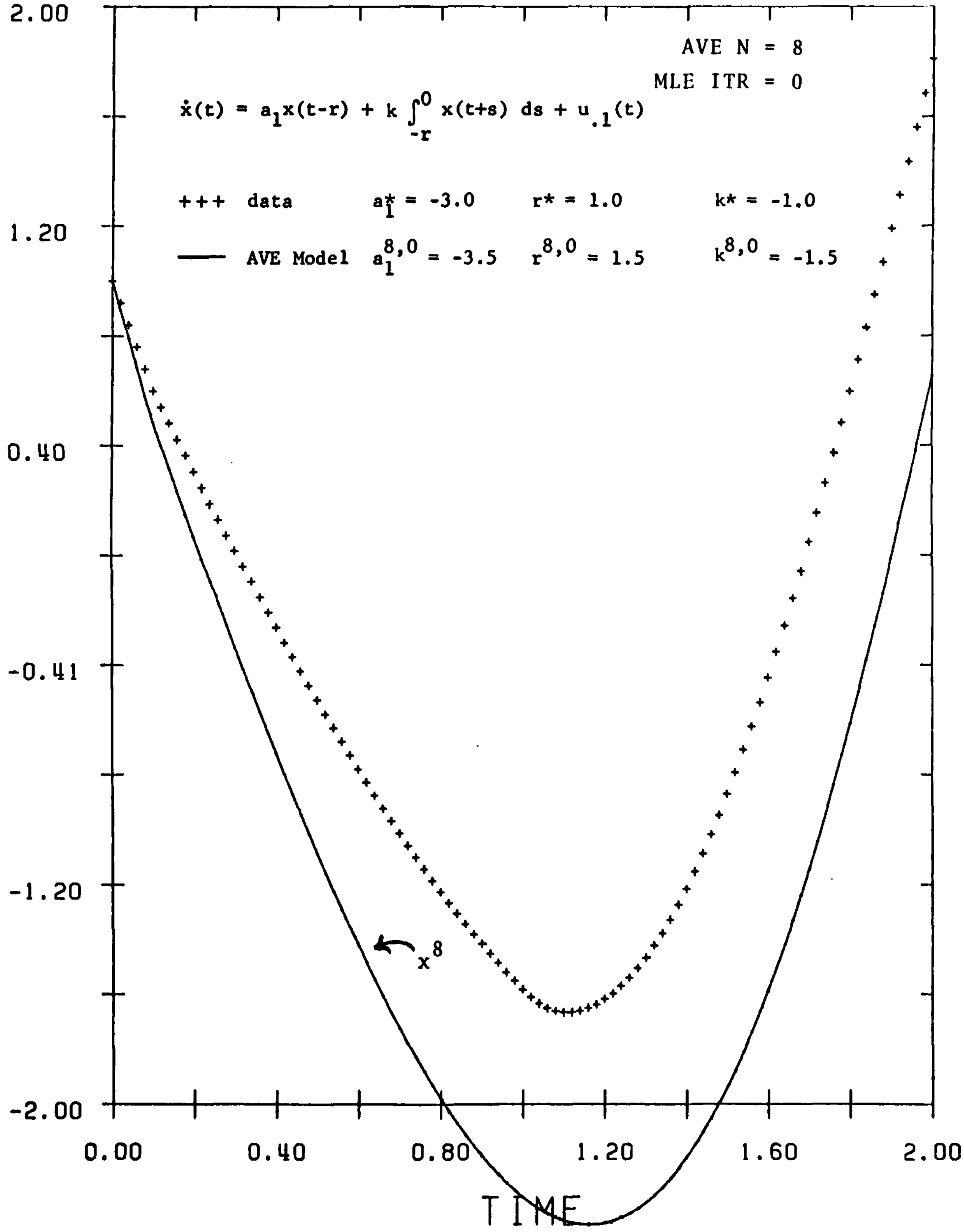

AVE N = 8
MLE ITR = 0
$\dot{x}(t) = a_1 x(t-r) + k \int_{-r}^{0} x(t+s)\, ds + u_{.1}(t)$
+++ data $a_1^* = -3.0$ $r^* = 1.0$ $k^* = -1.0$
AVE Model $a_1^{8,0} = -3.5$ $r^{8,0} = 1.5$ $k^{8,0} = -1.5$
x^8
2.00
1.20
0.40
-0.41
-1.20
-2.00
0.00
0.40
0.80
1.20
1.60
2.00
TIME

FIGURE 7.2.1

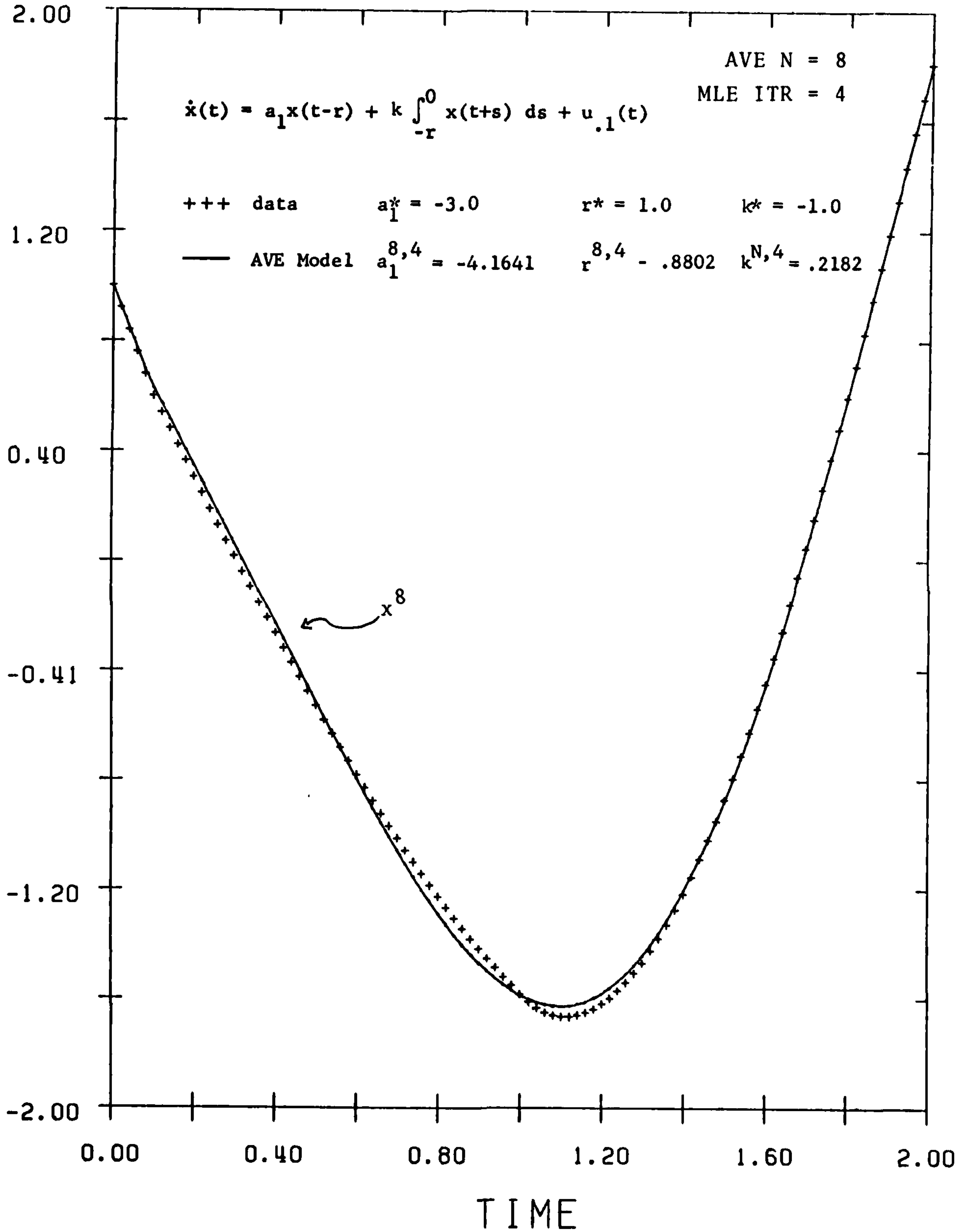
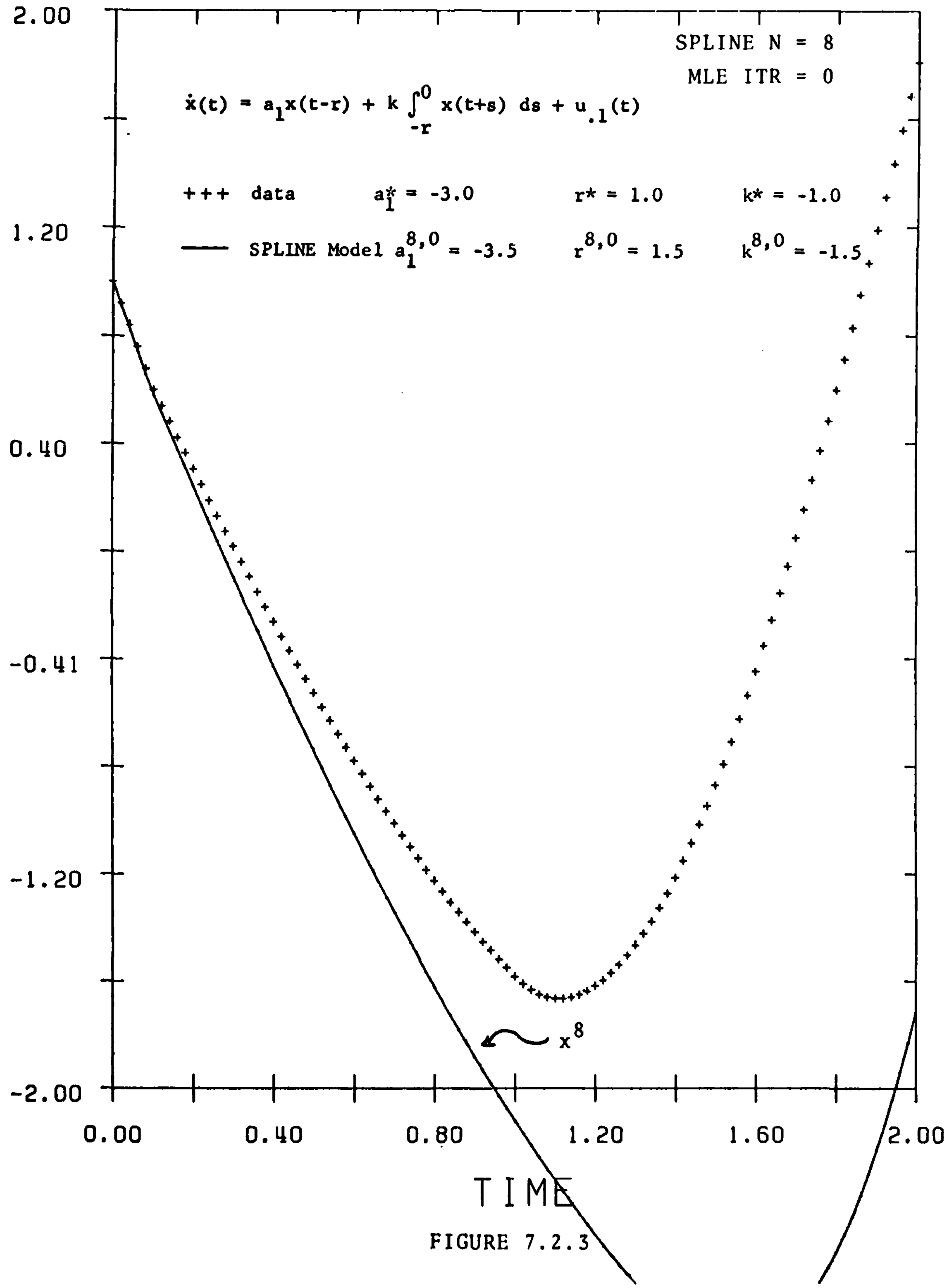
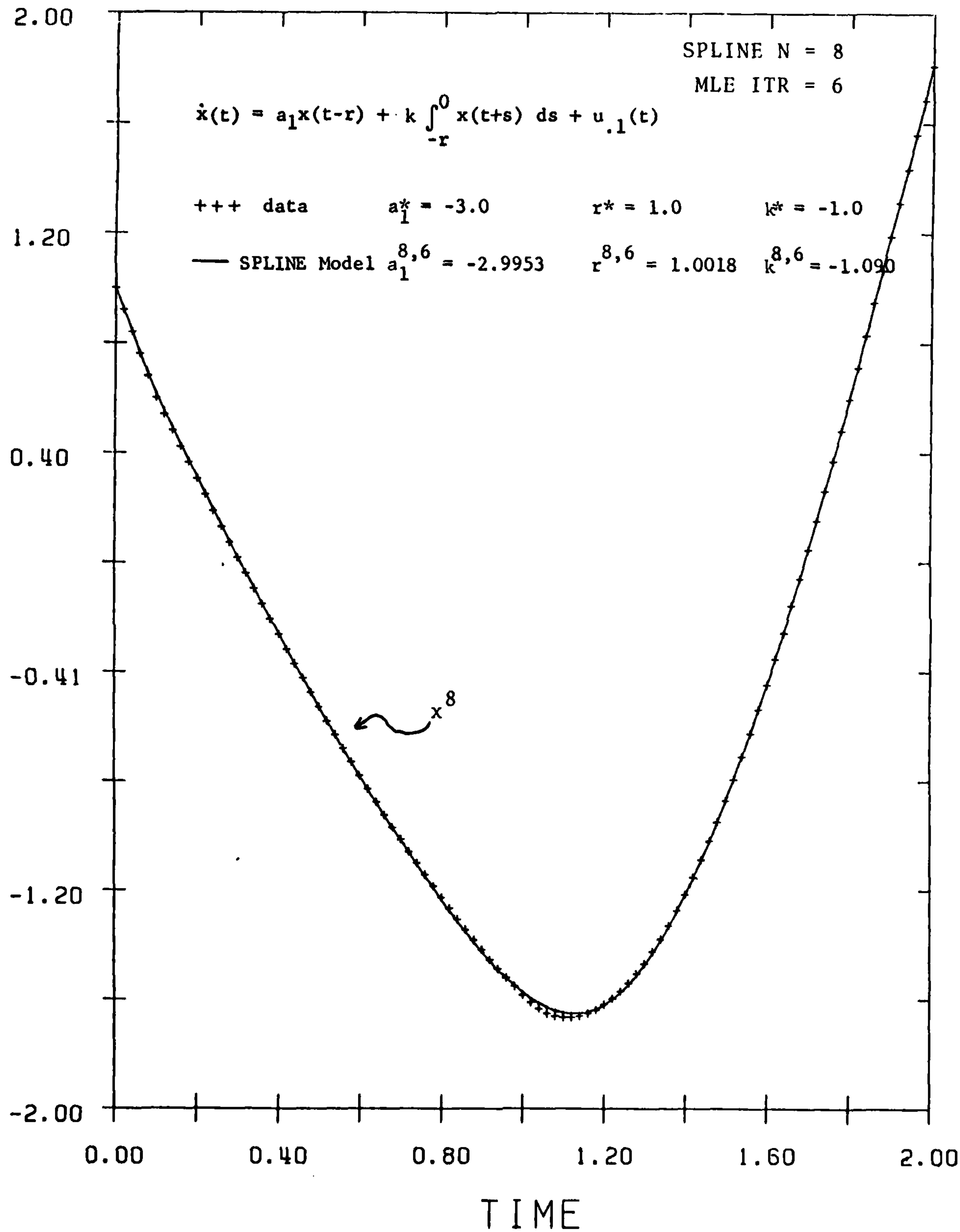

AVE N = 8
MLE ITR = 4
$\dot{x}(t) = a_1 x(t-r) + k \int_{-r}^{0} x(t+s)\, ds + u_{.1}(t)$
+++ data $a_1^* = -3.0$ $r^* = 1.0$ $k^* = -1.0$
AVE Model $a_1^{8,4} = -4.1641$ $r^{8,4} - .8802$ $k^{N,4} = .2182$
x^8
2.00
1.20
0.40
-0.41
-1.20
-2.00
0.00
0.40
0.80
1.20
1.60
2.00
TIME

FIGURE 7.2.2



2.00
1.20
0.40
-0.41
-1.20
-2.00
SPLINE N = 8
MLE ITR = 0
$\dot{x}(t) = a_1 x(t-r) + k \int_{-r}^{0} x(t+s)\,ds + u_{.1}(t)$
+++ data $a_1^* = -3.0$ $r^* = 1.0$ $k^* = -1.0$
SPLINE Model $a_1^{8,0} = -3.5$ $r^{8,0} = 1.5$ $k^{8,0} = -1.5$
x^8
0.00
0.40
0.80
1.20
1.60
2.00
TIME

FIGURE 7.2.3



SPLINE N = 8
MLE ITR = 6
$\dot{x}(t) = a_1 x(t-r) + k\int_{-r}^{0} x(t+s)\,ds + u_{.1}(t)$
+++ data $a_1^* = -3.0$ $r^* = 1.0$ $k^* = -1.0$
SPLINE Model $a_1^{8,6} = -2.9953$ $r^{8,6} = 1.0018$ $k^{8,6} = -1.090$
x^8
2.00
1.20
0.40
-0.41
-1.20
-2.00
0.00
0.40
0.80
1.20
1.60
2.00
TIME

FIGURE 7.2.4

Example 7.3

In our final example we consider an oscillator with retarded damping and retarded restoring forces. We seek to estimate the coefficients of the delayed terms and the time delay itself. The system is governed by the equation

$$\ddot{x}(t) + 16x(t) + a_0\dot{x}(t-r) + a_1 x(t-r) = u_{.1}(t),$$

with initial data

$$x_0(s) \equiv 1, \qquad \dot{x}_0(s) \equiv 0, \qquad -r \leq s \leq 0,$$

and output

$$y(t) = x(t).$$

This second-order equation is equivalent to the two-dimensional system

$$\frac{d}{dt}\begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -16 & 0 \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ -a_1 & -a_0 \end{bmatrix} \begin{bmatrix} x_1(t-r) \\ x_2(t-r) \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u_{.1}(t),$$

with initial condition

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}_0 (s) \equiv \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \qquad -r \leq s \leq 0,$$

and output

$$y(t) = [1\ 0] \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} .$$

The true parameters to be estimated are $a_0^* = 10.0$, $a_1^* = -10.0$ and $r^* = 1.0$. Start-up values for each run were

$$a_0^{N,0} = 11.0, \qquad a_1^{N,0} = -9.0, \qquad r^{N,0} = 1.2.$$

Convergence results for this example are summarized in Tables 7.3.1 and 7.3.2. At N = 16 the relative ℓ_1 error $(|e_N|/|\gamma^*|)$ for AVE is approximately 3.5%, while the N = 16 SPLINE scheme produced a relative ℓ_1 error of less than 1%.

Figures 7.3.1 and 7.3.2 show the N = 4 converged data fits for AVE and SPLINE, respectively. For $N \geq 8$, the data fits are nearly perfect and are not shown.

AVE				
N	$\bar{a}_0^N$	$\bar{a}_1^N$	$\bar{r}^N$	$\lvert e_N \rvert$
2	did not converge			
4	54.5124	-9.1876	2.4190	46.7439
8	19.4941	-9.4927	1.3506	10.3520
16	10.6433	-9.9089	.9998	.7346
γ^* =	10.0000	-10.0000	1.0000	

Table 7.3.1

SPLINE				
N	$\bar{a}_0^N$	$\bar{a}^N$	$\bar{r}^N$	$\lvert e_N \rvert$
2	9.2585	-10.5360	1.0908	1.3683
4	10.0927	-10.0619	1.0076	.1622
8	9.9724	-10.0177	1.0010	.0463
16	9.9811	-10.0108	1.0017	.0314
γ^* =	10.0000	-10.0000	1.0000	

Table 7.3.2

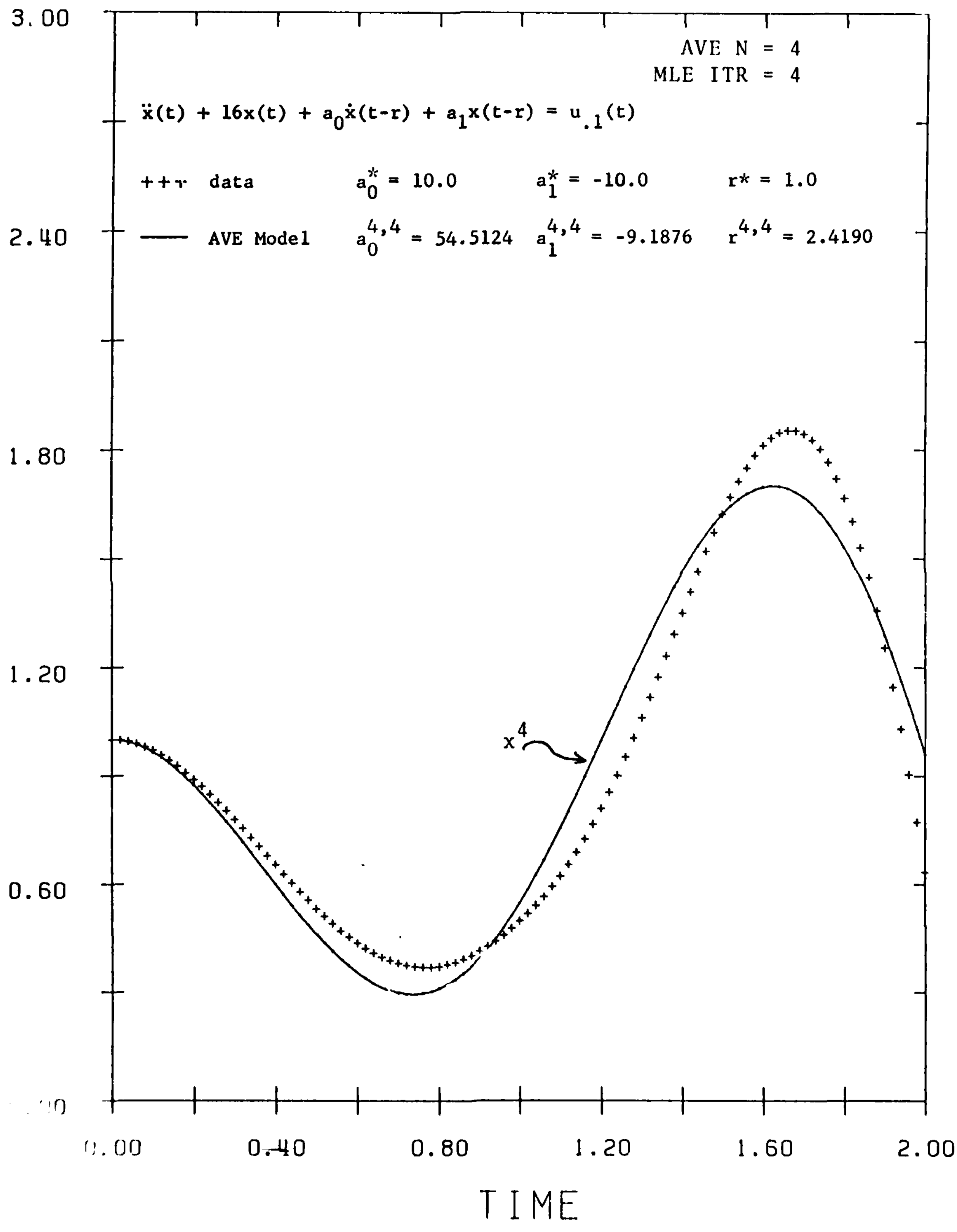
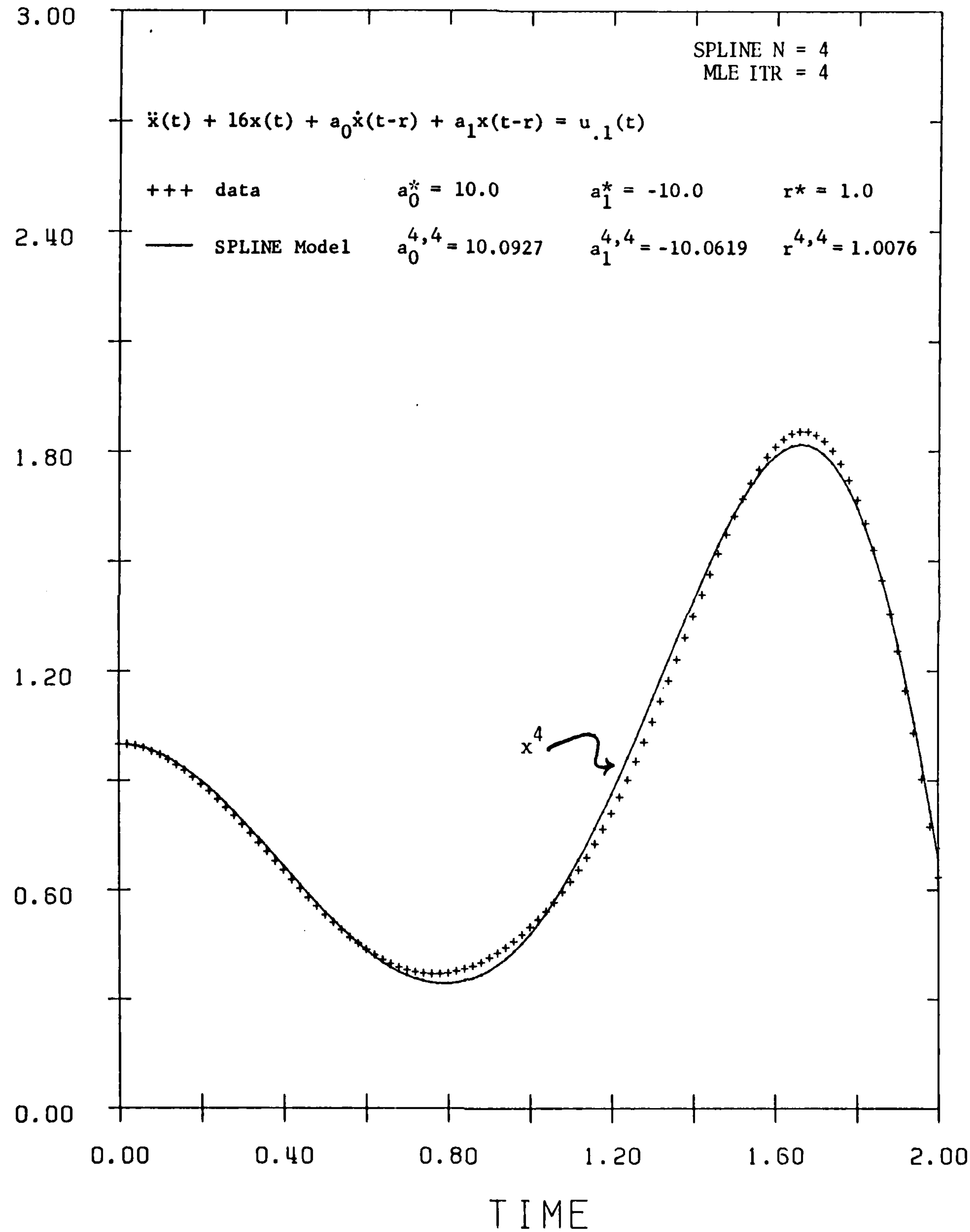

3.00
2.40
1.80
1.20
0.60
.00
AVE N = 4
MLE ITR = 4
$\ddot{x}(t) + 16x(t) + a_0\dot{x}(t-r) + a_1x(t-r) = u_{.1}(t)$
+++ data $a_0^* = 10.0$ $a_1^* = -10.0$ $r^* = 1.0$
AVE Model $a_0^{4,4} = 54.5124$ $a_1^{4,4} = -9.1876$ $r^{4,4} = 2.4190$
x^4
0.00
0.40
0.80
1.20
1.60
2.00
TIME

FIGURE 7.3.1



SPLINE N = 4
MLE ITR = 4
$\ddot{x}(t) + 16x(t) + a_0\dot{x}(t-r) + a_1x(t-r) = u_{.1}(t)$
+++ data $a_0^* = 10.0$ $a_1^* = -10.0$ $r^* = 1.0$
SPLINE Model $a_0^{4,4} = 10.0927$ $a_1^{4,4} = -10.0619$ $r^{4,4} = 1.0076$
x^4
3.00
2.40
1.80
1.20
0.60
0.00
0.00
0.40
0.80
1.20
1.60
2.00
TIME

FIGURE 7.3.2

References

[1] K.J. Åström and P. Eykhoff, System identification - a survey, Automatica 7 (1971), 123-162.

[2] A.V. Balakrishnan, Active control of airfoils in unsteady aerodynamics, Appl. Math. Opt. 4 (1978), 171-195.

[3] H.T. Banks, Approximation of nonlinear functional differential equation control systems, J. Optimization Theory Appl. 29 (1979), 383-408.

[4] H.T. Banks and J.A. Burns, An abstract framework for approximate solutions to optimal control problems governed by hereditary systems, in Proceedings, International Conference on Differential Equations (Univ. So. Calif., Sept., 1974), H.A. Antosiewicz, ed., Academic Press, New York, 1975, pp. 10-25.

[5] H.T. Banks and J.A. Burns, Hereditary control problems: numerical methods based on averaging approximations, SIAM J. Control and Optimization, 16 (1978), 169-208.

[6] H.T. Banks and J.A. Burns, Approximation techniques for control systems with delays, Proc. Int'l Conference on Methods of Mathematical Programming, 1977, Polish Scientific Publishers, Warsaw, to appear.

[7] H.T. Banks, J.A. Burns and E.M. Cliff, Spline-based approximation methods for control and identification of hereditary systems, in International Symposium on Systems Optimization and Analysis, A. Bensoussan and J.L. Lions, eds., Lecture Notes in Control and Info. Sci., Vol. 14, Springer, Heidelberg, 1979, pp. 314-320.

[8] H.T. Banks, J.A. Burns and E.M. Cliff, A comparison of numerical methods for identification and optimization problems involving control systems with delays, Brown Univ. LCDS Tech. Rep. 79-7, 1979, Providence, R.I.

[9] H.T. Banks, J.A. Burns, E.M. Cliff and P.R. Thrift, Numerical solutions of hereditary control problems via an approximation technique, Brown Univ. LCDS Tech. Rep. 75-6, 1975, Providence, R.I.

[10] H.T. Banks and F. Kappel, Spline approximations for functional differential equations, J. Diff. Eq. 34 (1979), 496-522.

[11] J.A. Burns and E.M. Cliff, On the formulation of some distributed system parameter identification problems, Proc. AIAA Symp. on Dynamics and Control of Large Flexible Spacecraft, 1977, pp. 87-105.

[12] J.A. Burns and E.M. Cliff, Methods for approximating solutions to linear hereditary quadratic optimal control problems, IEEE Trans. Automatic Control 23 (1978), 21-36.

[13] E.M. Cliff and J.A. Burns, Parameter identification for linear hereditary systems via an approximation technique, VPISU College of Engr. Tech. Rep., March, 1978; to appear in Proc. Workshop on the Linkage Between Applied Mathematics and Industry (Monterey, 1978).

[14] L.E. El'sgol'ts, Introduction to the Theory of Differential Equations with Deviating Arguments, Holden-Day, San Francisco, 1966.

[15] G. Gellf and J. Henry, Experimental and theoretical study of diffusion, convection and reaction phenomena for immobilized enzyme systems, in Analysis and Control of Immobilized Enzyme Systems, D. Thomas and J.P. Kernevez, eds., North Holland/ American Elsevier, New York, 1976, pp. 253-274.

[16] J. Henry, Contrôle d'un réacteur enzymatique à l'aide de modèles à paramètres distribués: quelques problèmes de contrôlabilité de systèmes paraboliques, Thèses d'État, Université Paris VI, 1978.

[17] T. Kato, Perturbation Theory for Linear Operators, Springer Verlag, New York, 1966.

[18] K. Kunisch, Approximation of optimal control problems for nonlinear hereditary systems of neutral type, to appear in Ber. math.-stat. Sekt. Forschungszentrum.

[19] T.G. Kurtz, Extensions of Trotter's operator semigroup approximation theorem, J. Functional Anal. 3 (1969), 354-375.

[20] P.D. Lax and R.D. Richtmyer, Survey of the stability of linear finite difference equations, Comm. Pure Appl. Math. 9 (1956), 267-293.

[21] M.Z. Nashed, Generalized inverses, normal solvability, and iteration for singular operator equations, in Nonlinear Functional Analysis and Applications, L.B. Rall, ed., Academic Press, New York, 1971, pp. 311-359.

[22] A. Pazy, Semigroups of Linear Operators and Applications to Partial Differential Equations, Math. Dept. Lecture Notes, Vol. 10, Univ. Maryland, College Park, 1974.

[23] A.P. Sage and J.L. Melsa, System Identification, Academic Press, New York, 1971.

[24] M.H. Schultz, Spline Analysis, Prentice-Hall, Englewood Cliffs, N.J., 1973.

[25] M.H. Schultz and R.S. Varga, L-splines, Num. Math. 10 (1967), 345-369.

[26] H.F. Trotter, Approximations of semigroups of operators, Pacific J. Math. 8 (1958), 887-919.

[27] H.F. Trotter, Approximation and perturbation of semigroups, in Linear Operators and Approximation II, P.L. Butzer and B. Sz.-Nagy, eds., Birkhaüser Verlag, Basel, 1974, pp. 3-21.

[28] T. von Karman and J.M. Burgers, General aerodynamic theory - perfect fluids, Vol. II, Aerodynamic Theory, W.F. Durand, ed., Dover, 1963.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER AFOSR-TR- 80-0403	2. GOVT ACCESSION NO. AD-A086 929	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) PARAMETER ESTIMATION AND IDENTIFICATION FOR SYSTEMS WITH DELAYS		5. TYPE OF REPORT & PERIOD COVERED Interim
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) H.T. BANKS, J.A. BURNS AND E.M. CLIFF		8. CONTRACT OR GRANT NUMBER(s) AFOSR 76-3092
9. PERFORMING ORGANIZATION NAME AND ADDRESS AIR DIVISION OF APPLIED MATHEMATICS BROWN UNIVERSITY PROVIDENCE, RHODE ISLAND 02912		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS 61102F 2304/A1
11. CONTROLLING OFFICE NAME AND ADDRESS AIR FORCE OFFICE OF SCIENTIFIC RESEARCH/NM BOLLING AIR FORCE BASE WASHINGTON, D.C. 20332		12. REPORT DATE November 1979
		13. NUMBER OF PAGES 87
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		15. SECURITY CLASS. (of this report) UNCLASSIFIED
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

Parameter identification problems for delay systems motivated by examples from aerodynamics and biochemistry are considered. The problem of estimation of the delays is included. Using approximation sehemes is developed and two specific cases ("averaging" and "spline" methods) are shown to be included in this treatment. Convergence results, error estimates, and a sample of numerical findings are given.

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE UNCLASSIFIED

ILMED

8